THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXXII November, 1941 Number 8

THE MEASURED INTERESTS OF ADOLESCENT GIRLS*

LEONA E. TYLER

University of Oregon

Measurement of the interests of adolescents has both practical and theoretical importance. The practical need arises from the fact that in the early high-school years the individual must make some tentative vocational plan, since he must choose a course of study at this time. The theoretical value lies in the fact that interests, as we are measuring them with our scales, are an integral and very important part of the complex of traits we call the personality. We need to know at what age they become noticeably differentiated and what their relationships are to other personality qualities and to environmental factors.

Most of the previous research on interest measurement has been based on the responses of adults, but there have been a few studies in which high-school students served as subjects. These give some evidence that clear-cut interest patterns may emerge in early adolescence. Garretson[7] constructed an interest blank for high-school boys which differentiated between students in Technical, Commercial, and Academic high schools. Carter, Pyles, and Bretnall[4] administered the Strong Test to one hundred thirty-three high-school boys. When scores on the different occupational scales were subjected to factor analysis, three factors and a possible fourth emerged. These were identifiable with those which Strong[13] had previously found for adults, although their relative importance seemed to be different. In a later study, Carter and Jones[3] found that a majority of the high-school

* This paper is a summary of research done under the direction of D. G. Paterson at the University of Minnesota for a M. S. thesis, *An Interest Test for Girls*, 1939, and a Ph. D. thesis, *Factors Conditioning the Development of Interests in Adolescent Girls*, 1940. Both theses are on file in the University of Minnesota library.

XUM

boys and girls who had made occupational choices received scores higher than *C* on the Strong scales most closely related to their choices. These findings seem to indicate that considerable vocational interest differentiation has already taken place by the time boys and girls reach the tenth grade.

Results from a number of investigations agree in showing that sex differences in interests characterize both adolescents and adults. Strong[14] showed that there were similar sex interest differences in high-school students, college students, and adults. Carter and Strong[5] and Finch and Odoroff[6] corroborated these findings for different groups of high-school students. Studies of stated vocational choices,[1,9,11] problems and areas of interest,[16] and subjects of conversation,[12,2] all give the same sort of results. Thus it would seem to be advisable to separate the sexes in any study of adolescent interests.

The present research is based on results obtained with a test constructed by the author, *The Minnesota Interest Test for Girls*. Many items were tried out with eighth-grade girls, and all those which more than twenty-five per cent of the unselected group failed to understand were eliminated. Others to which almost all subjects responded in the same way were also eliminated, since the purpose of the test was to make interest differentiations. In one preliminary study, the occupational items were rephrased, using a brief explanation of the type of work involved in each job in place of the job label. In randomly selected groups of high-school girls, responses to job label, job descriptions, and combinations of the two were compared to see whether or not the amount of prestige the different names carried was the factor determining the responses. There was no evidence that this was the case. The test as finally standardized consisted of one hundred ninety-nine items, including names of occupations, home, school and community activities, school subjects, and magazines. To each of these items the subject was asked to register an *L*, *I*, or *D* response. Scoring keys were based on the responses of seniors in three St. Paul high schools registered in the courses most commonly offered for girls, College Preparatory, Commercial, and General. Only individuals receiving at least average grades were included in the criterion groups. Items which differentiated College Preparatory girls from all the others were included in the College Preparatory (abbreviated C P) key. Items which differentiated Commercial girls from all the others were included in the Commercial key. No special key was constructed for the General group, because a pre-

liminary study showed it to be very heterogeneous, including both girls disqualified for academic work by low intelligence, and girls of high ability who were planning to enter some specialized but non-academic field such as art or music.

The validity of the scoring scales was checked by giving the test to the entire twelfth grade (one hundred thirty-three cases) and the entire eighth grade (one hundred twenty-six cases) in Muskegon Heights, Michigan. For the twelfth grade all differences between means of the three different curriculum groups were unquestionably significant, critical ratios running from 3.2 to 7.1. Overlapping was not extensive. Only four per cent of the Commercial students and eighteen per cent of the General students exceeded the mean of the College Preparatory group on the C P key. Only seven per cent of the College Preparatory girls and twenty per cent of the General girls exceeded the mean of the Commercial group on the Commercial key. The eighth-grade results could not be analyzed in this way since a large proportion of these subjects had not yet chosen their high-school course. However, means of the total group on the two scales were not significantly different from those of the total twelfth-grade group, and the distributions were similar. The reliability of the C P scale was .92 and that of the Commercial scale was .85 when a split-half correlation corrected by the Spearman-Brown formula was calculated for seventy-five twelfth-grade papers. Subsequent reliability calculations based on other groups have given similar results. Subsequent work in three other St. Paul high schools and in a Catholic and a non-Catholic private school have borne out the differentiations between curriculum groups. In general, it may be said that if critical scores for the two scales are set at suitable levels, more than three-fourths of the high-school girls in a typical class will be correctly assigned to the course of study they have chosen.

Satisfactory as this differentiation was, relationships of the scores to other factors, such as intelligence and fathers' occupations, made it appear desirable to find out on just what basis the differentiation was being made, since counselors using such a test in the diagnosis of individuals need to know exactly what the scores mean. It was also hoped that a detailed study of the interests and attitudes of high-school girls might contribute some valuable data to the understanding of women's place in modern society. Furthermore, the theoretical question of the nature and organization of personality traits was directly or indirectly involved in the study.

XUM

Relationships chosen for study on the basis of preliminary exploratory work were those between the interest scales and socio-economic status, parents' education, intelligence, attitude toward Feminism, and masculinity-femininity. The socio-economic classification was a four-step scale based on the Goodenough classification[8] supplemented by information obtained from responses to four home situation items. The educational classification was also a four-step scale based on the average of the last grades reached by the two parents. Intelligence was measured by the A C E Psychological Examination (1937 form). *A Belief Pattern Scale for Measuring Attitudes toward Feminism*, Form *A* by Clifford Kirkpatrick, and the *Attitude-Interest Analysis Test* by Terman and Miles (usually referred to simply as the M F test) completed the list of measures used. To begin with, the C P and Commercial scales of the Minnesota Interest Test were analyzed separately, but when the correlation between the two was found to be −.87, use of the Commercial scale was discontinued, since relationships found to hold for one were practically certain to hold in an opposite direction for the other. Research published by the authors of the various tests and a special reliability study of fifty cases selected at random from the present group of subjects indicated that all of them had reliabilities high enough so that correlations based on their scores would not be greatly attenuated.

The subjects were three hundred fifty-three twelfth-grade girls including practically the entire senior classes from Humboldt, Harding, Mechanic Arts, and Derham Hall high schools in St. Paul, and Northrop Collegiate School in Minneapolis. After eliminating from the principal sample the Northrop Collegiate subjects, who had taken the 1939 A C E test which could not be directly equated to the 1937 form, and all cases for which data were incomplete, two hundred sixty-seven cases remained. Various comparisons indicated that this sample was fairly typical of the Twin Cities high-school population in the percentage registered for each of the three courses of study, the percentage of private-school students, socio-economic status of families, and intelligence.

The first series of analyses was designed to furnish an accurate estimate of the extent of the relationships existing between these variables and to establish their statistical significance. While the correlational procedure could be used to ascertain relationships between tests and scales, it was not practical to handle the relationships of tests to the non-quantitative variables, socio-economic status

and educational level of parents, this way. For these relationships the analysis of variance procedure was adopted. It involves making two estimates of the population variance, one based on the means of the occupational or educational sub-groups, and one based on the variance within the groups. The ratio of these two estimated variances can then be tested for statistical significance using either Fisher's z test or Snedecor's F test. If the variance ratio is significant at the one per cent level, we are justified in concluding that the differences between subgroups are too large to be accounted for by sampling variation alone and that the scores must be significantly related to the criterion of classification. The direction of this relationship can be ascertained by examining the means of the subgroups. Its extent can be estimated by using the sums of squares already calculated for the analysis of variance test in the calculation of correlation ratios. In the present instance, results of the analyses can be summarized as follows:

(1) C P score was shown to be significantly and positively related to socio-economic level; $\eta = .37$

(2) Commercial score was significantly and negatively related to socio-economic level; $\eta = .32$

(3) C P score was significantly and positively related to educational level of parents; $\eta = .34$

It can be seen that the relationships between interest scores and level of home background, although they are statistically significant, are not high. It might be thought that relationship to a third common factor such as intelligence might account for the differences between groups. To test this hypothesis, an analysis of covariance was made for each of the three other variables for which quantitative scores had been obtained. This involves subtracting from each estimated variance the part that can be attributed to regression on the third factor in question, and then testing the ratios of the corrected variances for statistical significance. The results in this case furnished no support for the hypothesis. Whether the variances were adjusted for regression on A C E, attitude toward Feminism, or M F score, the differentiations between groups were still significant in every case. No one of these factors can be said to account for the relationship between interest score and level of home background.

Correlations between test scores are shown in Table I.

In general, it is evident that C P score correlates positively with all other scores except Commercial, and the Commercial score negatively

with all the others. While all except one of the correlations involving interest scores are statistically significant, they are too low to be very significant practically or theoretically. This fact set the problem for the next series of analyses, an attempt to account for the correlations in terms of psychologically meaningful relationships between the traits involved.

TABLE I.—CORRELATIONS BETWEEN FIVE TESTS USED IN THIS STUDY. PRINCIPAL SAMPLE OF TWO HUNDRED SIXTY-SEVEN CASES

Test	C P	Commercial	A C E	M F	Kirkpatrick
Interest test C P score	.94	−.85	.30	.20	.16
Interest test commercial score	...	.91	−.19	−.26	−.04*
A C E	...	.....	.....	.20	.11*
Terman-Miles M F	...	.....	.....	.78	.18
Kirkpatrick attitude toward feminism	...	.....	.....	.....	.82

* Not significant at one per cent level. Fisher's t test.

Since the M F test consists of seven subtests, intercorrelations between which are rather low, the first step was to ascertain whether or not any one of these was more closely related to C P score than was the total M F score. Exercises 1, 2, and 7 were eliminated at the outset because their reliability, measured by Jackson's test of sensitivity,[10] was so low that there was more than a five per cent probability that differences between individuals were no greater than the standard error of measurement. Of the correlations between C P score and the other four subtests, only those with Exercise 4, Emotional and Ethical Response, $r = .26$, and Exercise 6, Personalities and Opinions, $r = .15$, were statistically significant. These two exercises were subjected to an item analysis in which the responses of the fifty persons who had received the highest C P scores were tabulated and compared with the responses of the fifty persons who had received the lowest C P scores. Out of the six groups of items included in Exercise 4, only the two groups, headed "Fear" and "Wickedness," showed a clear trend and a large number of significant differences. The striking fact was that there was a marked tendency for the high C P group to use the extreme emotional reactions and condemnations less often and the moderate and more tolerant ones more often. Since Terman and Miles have shown that this is a masculine tendency, it seems

reasonable to attribute a large part of the apparent correlation between C P interests and masculinity to this correspondence. No significant results emerged from the item analysis of Exercise 6 except a tendency for the high C P group to register some reaction toward each personality named rather than to omit the item.

The Kirkpatrick scale was then subjected to the same sort of item analysis. There were ten of its eighty items for which the difference between proportions was significant at the one per cent level; seven for which it was significant at the five per cent level; and eleven for which it was significant at the ten per cent level. The trait that seemed to be involved in most of these significant items might be labeled "desire for independence" since they were concerned largely with economic equality, legal rights guaranteeing such equality, and the question of obedience to husbands. The same item analysis procedure was then applied to several items on the information blank that each subject had filled out. Whether or not parents were born in this country was not significant, but a significantly larger proportion of the high C P group was Protestant than Catholic. Responses on a graphic rating scale for registering strength of interest in several types of activity showed a significant tendency of the high C P group to specify strong interest in literature, art, and science.

Several factors having some psychological meaning had thus emerged from the item analysis. Girls receiving high C P scores differed from those receiving low scores in that they were less subject to extreme fears, more tolerant of conduct of which women often disapprove, more desirous of living independent lives on a basis of economic equality with men, and more strongly interested in the fields of literature, art, and science. Several questions were raised by these findings. Would the relationships between C P score and these other factors hold throughout the whole distribution of interest scores as well as at the extremes which had been used in the item analysis? That is, could we say, in general, that the higher a girl's C P score was the more likely she would be to show ethical tolerance and an independent attitude? How much of the variation in interest scores was based on these emotional and attitude factors which are not usually included under interests at all? Since variances are additive, and any total variance may be broken down into the part which may be attributed to regression on some related factor and the part which cannot be accounted for in this way, it was possible to get an approximate answer to these questions for the whole distribution if some

method could be found for measuring the special psychological factors under consideration. The M F Special and Kirkpatrick Special scores were worked out for this purpose. The M F Special score consisted of the scores made on only the Fear and Wickedness sections of Exercise 4, since the significant responses were in most cases the same as those already scored for masculinity. A new key for only the twenty-seven items which had shown significant differences between high and low groups gave the Kirkpatrick Special score. What these special scores served to do was to sort out of the rather heterogeneous collection of items brought together in each of these attitude tests the parts which seemed to be determined by psychological factors related to the C P interests. Correlation coefficients using the special scores were then calculated. They are summarized in Table II, which includes, for comparison purposes, the r with A C E score previously calculated.

TABLE II.—CORRELATIONS BETWEEN C P SCORE, A C E, AND M F AND KIRKPATRICK SPECIAL SCORES

Test	A C E	M F (special)	Kirkpatrick (special)
C P	.30	.35	.42
A C E	...	.39	.29
M F	...	...	.32

The most striking thing to be noted about these new correlations is that they are much higher than the ones using total scores. For the C P scale, the correlation with M F total score had been .20; with this part score it was .35. Similarly, the correlation with Kirkpatrick total had been .16; with the special score based on twenty-seven items it was .42. To get an estimate of the relative importance of the three factors, Beta coefficients were calculated. These are shown in Table III.

TABLE III.—BETA COEFFICIENTS FOR PARTIAL REGRESSION OF C P SCORE ON A C E, M F SPECIAL, AND KIRKPATRICK SPECIAL SCORES

12.34 (partial regression on A C E)	.1320
13.24 (partial regression on M F special)	.1992
14.23 (partial regression on Kirkpatrick special)	.3212

It appears that the factor represented by the Kirkpatrick Special score is most important, that represented by the M F Special score next in order, and intelligence least important.

The multiple correlation between C P score and all three of these factors taken together is .49. Fisher's *z* test shows this to be unquestionably significant for this number of cases. However, since R^2 equals .2449, it is evident that only about one-fourth of the variance has been accounted for by these factors. Consideration of the reliabilities of the special scales indicates that the importance of the Kirkpatrick Special score is being underestimated in these correlations, as the corrected split-half reliability of the 27-item scale is only .54. Relationships with the M F Special scale, which has a reliability of .86, are not thus attenuated. It is quite possible that a multiple R as high as .60 could be obtained by using a reliable measure of the factor which the Kirkpatrick Special scale represents. Slightly more than one-third of the variance might thus be accounted for. Other parts of the study suggest that some of the rest of it may be attributed to differences in regard to intellectual and cultural interests and "level of aspiration" based on family expectations, although no quantitative measure of these things is available.

It appears from this investigation that the C P score is not so much a measure of specific vocational interests as it is an indicator of certain basic attitudes. It corresponds more closely to Strong's Occupational Level scale than to those for specific vocations. Similarly, the high negative correlation probably indicates that the Commercial scale represents more non-professional than purely business interests. Since this seems to be the most fundamental differentiation in the interests of adolescent girls, it would be advisable to leave the College Preparatory group out of all comparisons in attempting to get differentiations between groups of girls in training for various non-professional occupations. Strong's generalization about the point of reference applies here also.[15] As long as groups of non-professional women are compared with professional women, they will all seem to be very much alike. When they are compared with each other, significant differences may appear. A revision of the Minnesota Interest Test based on these findings is in progress. The attempt will be made to construct non-professional interest keys for Office Work, Nursing, Housework, Beauty Culture, Selling, and possibly Elementary-school Teaching.

The item analysis of the Kirkpatrick scale indicated that both high and low C P groups tended to agree with most of the Feminist items. A surprising number of the girls, however, also agreed with anti-Feminist items which were directly opposed to the Feminist

XUM

items they had checked. This type of inconsistency was most marked in the low group, but was not entirely absent from the high group. For example, while seventy-seven per cent of them checked the Feminist item, "The relative amounts of time and energy to be devoted to household duties on the one hand and to a career on the other should be determined by personal desires and interests rather than by sex," sixty-six per cent of them also checked the anti-Feminist item, "A woman who continues to work outside the home after marriage is shirking her fundamental duty to home and children." The inconsistency suggests that many girls have accepted Feminist opinions, verbally at least, at the same time retaining traditional attitudes with respect to women's inherent limitations. Any striking improvement in women's economic position will probably have to be preceded by quite widespread changes in the attitudes of women themselves.

The theoretical importance of the research consists in its identification of a trait which cuts across the common boundaries of personality study, a trait involving attitudes, interests, and emotional responses. The line from complete irresponsibility to complete independence may represent a fundamental reference axis along which the personalities of women and girls may be measured. Some generalization rather than complete specificity of attitudes seems to be indicated, since these scores shown to be related are based on no common content. The method worked out for this problem, a combination of correlation and item analysis based on some reasonable hypothesis, is applicable in many other problems of accounting for low relationships in terms of factors having psychological meaning.

SUMMARY

(1) An interest test was constructed using items within the range of understanding and experience of eighth-graders, which differentiated between groups of high-school girls making satisfactory adjustments to the College Preparatory, Commercial, and General courses of study.

(2) Analysis of variance and covariance showed that interest scores were significantly related to the socio-economic and educational levels of the homes from which subjects came, and that these relationships could not be accounted for by dependence on any common factor, such as intelligence, which had been measured in this study.

(3) Low but significant positive correlations were obtained between C P interest score and A C E score, M F score, and Kirkpatrick score.

XUM

Low but significant negative correlations were obtained between Commercial interest score and A C E and M F scores. A high negative correlation between C P and Commercial scores was demonstrated.

(4) Item analysis indicated that the parts of the M F test most closely related to C P interests were the "Fear" and "Wickedness" sections of Exercise 4. The items from the Kirkpatrick test most closely related to C P interests were those dealing with the problem of economic independence for women.

(5) Special scales based on these differentiating sections and items produced much higher correlations with C P score than had total scores on the M F and Kirkpatrick scales, in spite of lower reliability because of decreased length.

(6) About one-fourth of the variance in C P score can be accounted for in terms of non-vocational factors such as objective emotional reactions, independent attitudes, and general intelligence. Were reliable tests for the first two of these characteristics available, it is likely that they would take an even more important place in the vocational interest picture.

(7) High-school girls, while accepting most of the Feminist contentions, also hold many markedly anti-Feminist opinions.

REFERENCES

1. Beeson, M. F. and Tope, R. E.: "A Study of Vocational Preferences of High-school Students," *Voc. Guid. Magazine*, Vol. VII, 1928, pp. 115–119.
2. Carlson, J. S., Cook, S. W., and Stromberg, E. L.: "Sex Differences in Conversations," *Jr. Appl. Psych.*, Vol. XX, 1936, pp. 727–735.
3. Carter, H. D. and Jones, M. C.: "Vocational Attitude Patterns in High-school Students," *Jr. Ed. Psych.*, Vol. XXIX, 1938, pp. 321–334.
4. Carter, H. D., Pyles, M. K., and Bretnall, E. P.: "A Comparative Study of Factors in Vocational Interest Scores of High-school Boys," *Jr. Ed. Psych.*, Vol. XXVI, 1935, pp. 81–98.
5. Carter, H. D. and Strong, E. K., Jr.: "Sex Differences in Occupational Interests of High-school Students," *Per. Jr.*, Vol. XII, 1933, pp. 166–175.
6. Finch, F. H. and Odoroff, M. E.: "Sex Differences in Vocational Interests," *Jr. Ed. Psych.*, Vol. XXX, 1939, pp. 151–156.
7. Garretson, O. K.: *Relationships between Expressed Preferences and Curricular Abilities of Ninth-grade Boys*, Teachers College Cont. to Ed., No. 396, 1930.
8. Goodenough, F. L.: *The Kuhlman-Binet Tests for Children of Pre-School Age*, Minneapolis: University of Minnesota Press, 1928.
9. Hurlock, E. B. and Jansing, C.: "The Vocational Attitudes of Boys and Girls of High-school Age," *Jr. Genet. Psych.*, Vol. XLIV, 1934, pp. 175–191.
10. Jackson, R. W.: "The Reliability of Mental Tests," *Br. Jr. of Psych.*, Vol. XXIX, 1939, pp. 267–287.

XUM

11. Lehman, H. C. and Witty, P. A.: "Sex Differences in Vocational Attitudes," *Jr. Appl. Psych.*, Vol. xx, 1936, pp. 576–585.
12. Stoke, S. M. and West, E. D.: "Sex Differences in Conversational Interests," *Jr. Soc. Psych.*, Vol. II, 1931, pp. 120–126.
13. Strong, E. K., Jr.: "Classification of Occupations by Interests," *Per. Jr.*, Vol. XII, 1934, pp. 301–313.
14. Strong, E. K., Jr.: "Interests of Men and Women," *Jr. Soc. Psych.*, Vol. VII, 1936, pp. 49–66.
15. Strong, E. K., Jr.: "Significance of the Point of Reference," *Psych. Bull.* Vol. XXXVI, 1939, p. 548.
16. Symonds, P. M.: "Life Problems and Interests of Adolescents," *Sch. and Soc.*, Vol. XLIII, 1936, pp. 751–752.
17. Tyler, L. E.: *Factors Conditioning the Development of Interests in Adolescent Girls*, Unpublished Ph. D. thesis on file in University of Minnesota Library, December, 1940.
18. Tyler, L. E.: *An Interest Test for Girls*, Unpublished master's thesis on file in University of Minnesota Library, June, 1939.

XUM

RELATIONSHIP BETWEEN MUSICAL TALENT AND PREFERENCES FOR DIFFERENT TYPES OF MUSIC

PAUL J. FAY AND WARREN C. MIDDLETON

DePauw University

Within recent years attention has been increasingly directed to preferences for different types of music. The problem has been given a somewhat more practical slant through the efforts of broadcasters to determine program preferences of radio listeners. Most radio surveys have primarily attempted to find out how many, and what kinds of people listen to particular programs. In some cases the results of the surveys have been analyzed with the purpose of determining what types of listeners prefer what types of programs. Music programs in radio surveys have usually been classified either as classical or popular.

The most important factor related to preference for classical music has been found to be the cultural (economic, social and educational) background of the listeners. According to Lazarsfeld,[1] "there is probably no question pertaining to radio listening which, by and large, is so sensitive to social differences as listening to serious music." This conclusion, of course, is consistent with the frequently expressed opinion of musicians and music critics that a liking for good music is a result of training in appreciation of, and frequent listening to, good music. Discriminative taste it is assumed, whether for good music, good art, or good wines, developes from frequent exposure to the good and abstinence from the bad. In our society prior to the broadcasting of operatic and symphonic music opportunity for the development of discriminative taste in music (as it still is, incidentally, in the case of art and wines) was limited to individuals in the upper economic levels. Carnegie Hall and the Metropolitan Opera House were beyond the means of most American citizens. In addition, "culture" (meaning a taste for, indulgence in, and the ability to converse entertainingly about classical music, art, literature and drama) was (or, should we say, *is*) a caste mark; as such, it was jealously guarded by the socially elite. What wonder, then, that children who grew up in socially and economically less fortunate homes were rarely encouraged to develop a taste for better music.

---

[1] Lazarsfeld, P. F.: *Radio and the Printed Page.* Duell, Sloane & Pearce, 1940, p. 22.

XUM

Radio, it may be argued, has changed all this. Anyone, whether in penthouse or hovel, may listen to the General Motors Symphony or to the Metropolitan Opera. If, instead, he turns the dial to the National Barn Dance, well, doesn't that indicate that his tastes are naturally depraved? Possibly. Or, possibly, some social traditions and habits are too deep-seated to change in the relatively short time that radio has been a dispenser of culture. Since radio broadcasting has scarcely yet emerged from the experimental stage, it seems unjustifiably dogmatic to conclude that it can never improve the musical taste of the general public.

There can be no disagreement about the present relationship between listening preferences and cultural level. For example, the data analyzed by the Coöperative Analysis of Broadcasting for the Columbia University Office of Radio Research show that twenty and three-tenths per cent of the radio listeners in the wealthiest of three economic levels listened to the Ford Hour; in the lowest of the three groups (*i.e.*, homes with incomes of two thousand dollars to three thousand dollars) only thirteen and seven-tenths per cent listened to the same program.[1] In another study sixty-two per cent of the lower and twenty-nine per cent of the higher income groups liked hill-billy music.[1] The Office of Radio Research also analyzed the data from a survey conducted by the Book-of-the-Month Club in which the Club members were classified into five levels according to their tastes for literature. Twenty-four per cent of the individuals in the highest level preferred classical music programs; the proportion dropped steadily to the lowest group in which seven per cent preferred classical music. Popular music, on the other hand, rose in favor from thirteen and five-tenths per cent in the highest group to fifteen and one-tenth per cent in the lowest group.[2]

Wiebe[3] unearthed another factor which influences musical preferences—a factor which is usually not included in "cultural background," yet one which is definitely a part of the cultural pattern. He studied the effect of radio plugging on high school and college students' ratings of popular songs. The results indicated that, whereas the liking for unplugged songs decreased within a month, two months after

---

[1] Lazarsfeld, P. F.: *Ibid.*, p. 22.

[2] Lazarsfeld, P. F.: *Ibid.*, pp. 35–41.

[3] Wiebe, G.: "The effect of radio plugging on students' opinions of popular songs." *Jour. Appl. Psychol.*, Vol. XXIV, 1940, pp. 721–727.

the original rating plugged songs were rated approximately the same.[1]

The purpose of the present experiment is to determine the relationship between musical talent and preferences for different types of music. We assume that musical talent is less a product of the environment than are wealth, social position, education and cultural tastes. We have attempted to use a finer classification than that employed in most radio surveys. After all, there are several kinds of classical music and at least two kinds of popular music. Some people prefer one type of classical music to another. Are these preferences entirely a result of training? And with identical cultural backgrounds would all individuals enjoy equally well the same types of music? As a matter of fact, we are quite sure that many penthouse dwellers prefer Benny Goodman to Gladys Swarthout.

The experiment was performed in the DePauw Laboratory for Research on the Psychological Problems of Radio. A detailed and illustrated description of the equipment and facilities of the Laboratory may be found in another article by the authors.[2]

The listeners, or judges, consisted of fifty-four students (twenty-seven men and twenty-seven women) in educational psychology classes. They were not, to be sure, a representative group of radio listeners. For the purpose of the experiment a representative, or heterogeneous, group of listeners would have been undesirable. By limiting the experimental group to college students we secured individuals who were fairly well equated with respect to cultural background.[3] Since all the listeners were prospective public-school teachers, the group was even more homogeneous than the student body

[1] Since this article went to press, Suchman has published a significant study on the contribution made by radio to musical education and on the radio listener's approach to music. See Suchman, Edward A.: "Invitation to Music." *In* Lazarsfeld, P. F., and Stanton, F. N. (ed.): *Radio Research* 1941. Duell, Sloan & Pearce, 1941, pp. 140–188.

[2] Fay, P. J., and Middleton, W. C.: "The DePauw Laboratory for Research on the Psychological Problems of Radio." *Amer. Jour. Psychol.*, Vol. LIV, 1941, pp. 571–575.

[3] The student body of DePauw University is much more homogeneous than the student bodies of most state or municipal universities. The great majority of the students come from upper middle class, urban, Middle Western homes. More than half their fathers belong to the professional and managerial groups of occupations. *Cf.* Fay, P. J., and Middleton, W. C.: "Certain factors related to liberal and conservative attitudes of college students: Sex, classification, fraternity membership, major subject." *Jour. Educ. Psychol.*, Vol. XXX, 1939, p. 379.

XUM

in general. Any differences, then, may be considered individual rather than group or cultural.

Two weeks prior to the experimental session three of the Seashore *Measures of Musical Talent* (Sense of Pitch, Sense of Rhythm and Sense of Time) were administered to the listeners. The adult percentile rank on each of the three tests was computed for each student.

At the beginning of the experiment each listener was furnished a mimeographed rating scale on which were listed musical selection numbers from 1 to 12. Each selection number was followed by the scale -3, -2, -1, 0, $+1$, $+2$, $+3$. All instructions were broadcast by the announcer over the public address system. After preliminary recorded music the announcer read the following instructions:

> You will hear several musical selections. As soon as each selection is finished, rate it for pleasantness or unpleasantness. In making each rating do not consider your ratings of previous selections; avoid any comparisons of the musical selections. Your judgment should be naïve and immediate; that is, not "thought out," but based simply upon the way the music appeals to you. Try to let the same number indicate the same degree of pleasantness or unpleasantness throughout the experiment.

The listeners were not told the purpose of the experiment or what types of music were included in the program.

The music, which was played over the public address system, consisted of twelve complete commercial recordings; two selections represented each of six different types of music. No two selections of the same type were played consecutively; the order of playing, as shown in Table I, was random. The classification of the selections was the consensus of three faculty members of the Music School of DePauw University. The selections classified into their respective types were the following:

Romantic Classical
- Tchaikovsky, *Marche Slav* (1st face).
- Wagner, *Flying Dutchman Overture* (1st face).

Old Classical
- Haydn, *Surprise*, First Movement.
- Beethoven, *Coriolan Overture* (1st face).

Modern Classical
- Ravel, *Alborado del Gracioso* (1st face).
- Rimski-Korsakov, *Capriccio Espagnol* (1st face).

XUM

Light Classical
Herbert, *March of the Toys.*
Herbert, *Naughty Marietta,* Intermezzo.
Swing[1]
Dorsey, *Stop, Look and Listen.*
Shaw, *My Heart Stood Still.*
Sweet
Whiteman, *Stardust.*
Scott, *Dinner Music.*

The mean ratings of the pleasantness of the different types of music are given in Table I. All ratings are positive; no type of music

TABLE I.—MEAN RATINGS OF PLEASANTNESS OF DIFFERENT TYPES OF MUSIC

Type of music	Orders[1]	Listeners' ratings of pleasantness[2]						Reliability coefficient[6]		
		Men[3]	Women[3]	D[4]	D/σ_D	Total[5] M	Total[5] σ	r	PE	Corrected r[7]
Light classical	6 and 9	+1.76	+2.61	W .85	3.70	+2.19	1.26	.60	.05	.75
Old classical	4 and 12	+1.39	+2.09	W .70	3.03	+1.74	1.26	.47	.07	.64
Romantic classical	1 and 11	+1.31	+2.17	W .86	3.48	+1.74	1.36	.46	.07	.63
Swing	2 and 10	+1.46	+ .65	M .81	3.13	+1.06	1.41	.54	.06	.70
Modern classical	3 and 7	+ .43	+1.06	W .63	2.26	+ .75	1.49	.39	.08	.56
Sweet	5 and 8	+ .35	+ .57	W .22	1.03	+ .24	1.12	.03	.09	.05
Mean		+1.12	+1.52	W .40	2.50	+1.32	1.46			
Total								.35	.03	.52

[1] The two orders in which each type of music was heard.
[2] The listeners rated each musical selection for pleasantness on a seven-point scale from −3 to +3.
[3] $N = 27$.
[4] The difference between the mean ratings of the sexes is preceded by an M, if the men gave a higher rating to the type of music; if the women gave it a higher rating, it is preceded by a W.
[5] $N = 54$.
[6] Correlation between ratings for the first selection heard and for the second selection heard within a given musical type.
[7] Corrected by the Spearman-Brown prophecy formula.

was given a mean unpleasant rating. The light classical numbers were quite definitely preferred; they had a mean rating of +2.19. The fact that light classical was rated higher than serious music is one indication

[1] At the time the study was made these were current popular songs.

that the listeners were rating sincerely; apparently, they were not unduly influenced by the possible opinion of the experimenters. There is considerable probability that the ratings for light, old and romantic classical music and for swing music represent measures of preference for the types of music as types; the ratings for sweet music, however, are apparently reactions to the particular selections played.

The reliability of the mean ratings was computed by correlating the ratings of the first selection of each type presented with the ratings of the second selection of the same type. Since only two different representatives of each type were rated, corrected r's ranging from .63 to .75 should not be considered unduly low. The mean ratings, then, of light, old and romantic classical music and of swing music are satisfactorily reliable. The reliability coefficient for sweet music, however, is practically zero; either the two selections of sweet music which were used are not truly representative of their type, or else there is no stable liking for sweet music as such. Casual observation and the results of Wiebe's study[1] would lead to the expectation that preferences for popular music are less stable than preferences for classical music. The ratings for modern classical music are fairly reliable (corrected $r = .56$); they are less reliable, however, than the ratings of other types of classical music. This finding also agrees with casual observation.

The fact that classical music is preferred to popular music is consistent with the results of the radio surveys. The listeners in the experiment would be classified in the higher cultural levels.

Women rate all types of classical music higher than the men do; the critical ratios (D/σ_D) range from 2.26 to 3.70. Men rate swing music higher; the critical ratio is 3.13. The women's mean rating for all twelve selections is .40 higher; the critical ratio is 2.50. We have no way of knowing whether or not these sex differences are typical of the general population, although they do agree with popular opinion. Women, particularly college women, are popularly believed to have more "cultural" interests than men.

The relationship between musical talent and preferences for the different types of music may be measured in various ways. One method is illustrated in Table II. The average rating of the two selections representing each type of music was computed for each listener. Some students rated one type of music higher than any of the other types; others rated two or three types of music equally high. If a

[1] Wiebe, G.: *Ibid.*

TABLE II.—MUSICAL TALENT PERCENTILES OF GROUPS PREFERRING DIFFERENT TYPES OF MUSIC

Type of music preferred by group	N[1]	Musical talent percentile[2]					
		Pitch		Rhythm		Time	
		M	σ	M	σ	M	σ
Romantic classical	17	64	24.90	82	23.70	69	28.90
Old classical	15	64	30.50	71	30.90	66	31.50
Modern classical	5	57	29.90	73	31.90	57	27.90
Light classical	30	49	26.10	61	36.30	46	30.60
Swing	14	32	24.00	49	29.40	46	30.20
Sweet	2	75	20.00	95	0	95	0

[1] Includes number of listeners whose highest mean rating was assigned to a given type of music. If a listener's highest rating resulted in a tie, his preference was recorded in all the types of music tied.

[2] Percentile for adults on Seashore *Measures of Musical Talent.*

listener rated one type higher than he did any of the other types, he was recorded with the group which he rated highest. If his highest rating was the same for two or more types of music, he was recorded with all of the tied types. As a result of this duplication of preferences the total of the "N" column is greater than the total number of listeners.

Since only five listeners rated modern classical music highest and only two rated sweet music highest, the data from these groups are not reliable. In fact, the two listeners who rated sweet music highest had the same mean rating for sweet music and for each of the four types of classical music; apparently, they liked equally well all kinds of music except swing.

As shown in Table III, the differences between the mean musical talent percentiles of some of the groups are statistically reliable. The critical ratio is greater than 3.00 for two differences in sense of pitch and for one difference in sense of rhythm. The critical ratio is between 2.00 and 3.00 for one difference in sense of pitch, for one in sense of rhythm, and for three in sense of time.

The results indicate that individuals who prefer romantic or old classical music are decidedly above the average of the adult population in the three measures of musical talent. Those who prefer light classical music are about average in sense of pitch and time, and slightly

above average in sense of rhythm. Lovers of swing, however, are slightly below average in sense of rhythm and time, and decidedly below average in sense of pitch. The popular idea that swing devotees have a strong sense of rhythm is not supported by the results of this experiment. As a matter of fact, in all three measures of musical talent they are decidedly inferior to those who prefer serious classical music. In two of the measures they are even inferior to those who prefer light classical music.

TABLE III.—DIFFERENCES BETWEEN MUSICAL TALENT PERCENTILES OF GROUPS PREFERRING DIFFERENT TYPES OF MUSIC

Groups compared	Pitch			Rhythm			Time		
	D[1]	σ_D	D/σ_D	D[1]	σ_D	D/σ_D	D[1]	σ_D	D/σ_D
Romantic classical Old classical	0	.....	....	11	9.84	1.12	3	10.74	.28
Romantic classical Light classical	15	7.69	1.95	21	8.77	2.39	23	8.98	2.56
Romantic classical Swing	32	8.82	3.63	33	9.74	3.39	23	10.69	2.15
Old classical Light classical	15	9.21	1.63	10	10.37	.96	20	9.88	2.02
Old classical Swing	32	10.16	3.15	22	11.20	1.96	20	11.46	1.75
Light classical Swing	17	7.99	2.13	12	12.47	.96	0		

[1] The difference between the mean percentiles of the two groups is printed in the same row as the name of the group having the higher mean.

Since the ratings of sweet music are unreliable, it may not be valid to draw any conclusions concerning preferences for popular music as a whole. The selections we have classified as swing, however, are probably more representative of popular music than are the sweet selections. The results do support the conclusion that preferences for classical or popular music are not entirely a product of the cultural background. We do not belittle, to be sure, the importance of the cultural factor. We merely suggest that the low degree of musical

talent[1] possessed by some individuals in the upper cultural levels may prevent their taking advantage of the opportunities offered them to develop a taste for classical music. On the other hand, many individuals in the lower social and economic levels may possess considerable musical talent and yet—from lack of opportunity—may never acquire a taste for classical music.

Different degrees of musical talent not only distinguish lovers of classical music from devotees of swing, but they also seem to be related to preferences for different types of classical music. Those who prefer romantic classical music are somewhat superior in all three measures of talent to those who prefer the other two types of serious music. None of the differences are statistically reliable, but they are consistent. For the purpose of distinguishing groups who prefer different types of music, differences in sense of pitch and rhythm are more significant than differences in sense of time. The respective means of the six differences listed in the "*D*" columns of Table III are: pitch, 18.50; rhythm, 18.17; time, 14.83.

The relationship between musical talent and preferences for the different types of music is also indicated in Table IV. These data

TABLE IV.—CORRELATION BETWEEN MUSICAL TALENT AND RATINGS OF PLEASANTNESS OF DIFFERENT TYPES OF MUSIC

Type of music	Pitch		Rhythm		Time		Pitch and rhythm		Pitch, rhythm and time	
	r	PE	*r*	PE	*r*	PE	R^1	PE	R^1	PE
Romantic classical	.46	.07	.43	.08	.21	.09	.50	.07	.505	.07
Old classical	.39	.08	.31	.09	.15	.09	.41	.08	.41	.08
Modern classical	.28	.09	.13	.09	.12	.09	.33	.09	.28	.09
Sweet	.24	.09	.10	.09	.26	.09	.22	.09	.47	.07
Light classical	.08	.09	.01	.10	.12	.09	.14	.09	.14	.09
Swing	−.16	.09	−.14	.09	−.13	.10	.22	.09	.35	.09

[1] Coefficient of multiple correlation.

corroborate those appearing in Tables II and III. One would not expect to find very high correlations between ratings of pleasantness of

[1] This study is not concerned with the problem of the extent to which musical talent is innate or acquired. Musical talent, however, does seem to be quite unrelated to the factor of cultural level. College students certainly vary much more in musical talent than they do in cultural environment.

different types of music and the musical talent scores of the raters; there are, unquestionably, several other variables. Yet the results do show that ratings of romantic, old and modern classical music—in that order—are positively related to the three musical talents measured. The chance coefficient of multiple correlation with three variables is .19.[1] The multiple correlation between ratings for romantic classical music and sense of pitch and rhythm is more than three PE's greater than the chance R. The corresponding coefficient for old classical music is almost three PE's greater than the chance R.

The correlation between musical talent and a liking for light classical music is practically zero. Preference for swing music has a small but consistent negative correlation with musical talent. The data seem to indicate that, other things being equal, individuals who prefer swing music are lacking in musical talent. Their preferences for particular musical selections are dependent upon factors other than the musical quality of the selections. Wiebe's[2] results tend to substantiate this conclusion. It is doubtful if radio plugging, or any similar extraneous factor, would affect many individuals' ratings of serious music.

The data in Table IV support the conclusion that sense of pitch and rhythm are more significant than sense of time in distinguishing preferences for the different types of music; the coefficients of multiple correlation for all three measures of musical talent are not reliably higher than those for pitch and rhythm alone.

From the standpoint of radio the results of this study would seem to support the hypothesis that persistent broadcasting of serious music may eventually increase the number of people who appreciate and listen to it. To the extent that the musical talents measured by the Seashore tests are innate they are distributed more or less equally among the members of different social and economic groups. Partly because appreciation of classical music has been a prerogative of a caste, listeners to broadcasts of serious music have to a great extent been individuals from the upper economic levels. But if the opportunity of listening to good music remains available to all at no cost, there is a possibility that the caste tradition will eventually disappear. In time, then, all individuals—irrespective of their economic status—who possess the requisite musical talent may acquire a discriminative musi-

[1] Derived from the formula $R = \frac{\sqrt{n-1}}{\sqrt{N}}$, in which n = the number of variables, and N = the number of cases.

[2] Wiebe, G.: *Ibid.*

cal taste. For individuals who are sufficiently gifted—but only for such individuals—frequent exposure to the good may lead to a preference for the good.

The conclusions from this study may be summarized as follows:

(1) College students' ratings of classical and swing music are fairly reliable.

(2) College students rate classical music higher than popular music.

(3) College women rate classical music higher than college men rate it; college men rate swing music higher than college women do.

(4) Individuals who prefer romantic classical music are slightly superior in sense of pitch and rhythm to those who prefer old and modern classical music; they are decidedly superior in musical talent to individuals who prefer light classical music.

(5) Individuals who prefer swing music are decidedly inferior in sense of pitch, rhythm and time to those who prefer classical music.

(6) Sense of pitch and rhythm are more significant than sense of time in distinguishing preferences for different types of music.

(7) Continued broadcasting of serious music may contribute to the improvement of the musical tastes of talented individuals in the lower economic levels.

XUM

THE PROGNOSTIC VALUE OF ENTRANCE TESTS IN A JUNIOR COLLEGE[1]

JOSEPHINE M. SMITH

Institute of Child Welfare, University of California

I. INTRODUCTION

The public junior college differs from the four-year-college, university, and private junior college in several important ways, not the least of which is the composition of its student body. While the latter three types of institution select their students according to scholastic standards and exclude those applicants who do not satisfy the requirements, the public junior college accepts all graduates of any high school, regardless of academic standing. Consequently entrance tests must subserve entirely different purposes in the public junior college than they do in the other colleges. The chief functions of such tests in this extension of the secondary-school system are to determine the type of student body, to predict the level of college achievement, and to guide students to the choice of subjects in which they are most likely to have success.

II. THE PROBLEM

The immediate problem is to discover which, if any, of the tests administered during the freshman year in a typical public junior college[2] serve as useful guides in predicting passing grades and to evaluate the predictive value of each. More specifically, the problem is to answer the following questions:

(1) What characteristics differentiate those students who complete the first college year with passing grades from those who do not?

(2) What characteristics differentiate those students who make good grades from those who make poor grades?

III. REVIEW OF THE LITERATURE

In 1934 David Segel published a review of studies up to that date. His summary shows the medians of correlation coefficients between

[1] Acknowledgment is gratefully made to Dr. H. E. Jones and Dr. H. S. Conrad of the Institute of Child Welfare, University of California, for reading and constructive criticism of this paper.

[2] The students whose records are analyzed here were registered in the Stockton Junior College, at Stockton, California.

single predictive measures and average college marks to be

General mental tests	.44
General achievement tests	.545
Tests of specific traits, aptitudes or achievements	.37
Average high school marks	.55

The direct relationships between combinations of predictive items and various criteria of college success vary from a correlation of .56 for a combination of Ohio Psychological examination and average high-school marks, to .81 for a combination of average high-school marks, Otis group intelligence test, and Cross English test. The combinations include batteries of from two to six different tests. After a summary of methods used, Segel says that "The problem of admission has to be met by each individual institution according to the kind of student body it wishes to have and its objectives of education."[1] The highest predictive value is that of the high-school average grades (.55), with general achievement tests ranking a close second (.545), general mental tests third (.44), and tests of specific traits or aptitudes last (.37).

Studies since Segel's review show the same relative values for the types of tests used, with batteries of tests ranking first. For example, Wagner[2] found the median correlation coefficients between criteria and average college marks to be

High school average	.56
Intelligence tests	.45
Batteries	.67
Regents' average	.64
Iowa Content	.49

Another significant study on this subject is that of Landry,[3] with the following correlations with college grades:

Adjusted mean secondary marks	.625
College Entrance Examination Board (Subject examinations)	.491
Coöperative tests	.573

IV. METHOD AND DATA OF THE PRESENT PROJECT

The present problem, as indicated, is to find which of the measures available at the time of registration of the freshman student are helpful in predicting the probable scholastic success of that student in the

[1] (13), p. 9.
[2] (15), (16).
[3] (9).

junior college. Our purpose differs from that of the studies cited above to the degree that the public junior college differs from the four-year institution. "Success" is here defined as achievement of a grade point average of one or above, since this G.P.A. is required for the Junior Certificate. The scholastic line between terminal and continuing students is drawn at this point.

The subjects were divided into two groups, those attending the college one year only and those attending for two years. Each group was subdivided into two, the "success" and the "failure" groups, and comparisons made between the means of the prognostic measures. Correlations between grades and the scores on the various prognostic measures were also determined.

V. DESCRIPTION OF THE GROUP

The subjects of the study were the two hundred sixty-seven students registering in the Fall of 1936 who remained at least one academic year. Measures on the group were available for high-school grade point average, Thurstone scores, Bell Emotional Adjustment scores, and Iowa Reading Test scores for comprehension and for rate. The students made records as shown below.

One term only	29	
Success		12
Failure		17
One year only	82	
Success		35
Failure		47
Two full years	185	
Success two years		117
Failure two years		60
Success first year, failure second		8

VI. RESULTS

A. Differentiation between Success and Failure Students.—Table I shows that the high-school grades, the Thurstone scores, and the Iowa Reading comprehension scores differentiate reliably between the success and failure groups of the first year; the Iowa Reading rate scores contribute somewhat.

Table II indicates that students who have G.P.A.'s of 1 or above for two full years make reliably higher scores than those whose G.P.A.'s are below 1 for one or two years, in high-school G.P.A., Thurstone score, Iowa Reading comprehension, and Iowa Reading rate scores.

XUM

TABLE I.—COMPARISON OF FIRST-YEAR SUCCESSES WITH FIRST-YEAR FAILURES

Measure	Success mean (n = 35)	Failure mean (n = 47)	D/σ_D
Thurstone	184.48	144.21	6.80
High-school G.P.A	2.97	2.30	11.25
Iowa comprehension	143.22	122.12	5.40
Iowa rate	28.29	25.96	2.58
Bell adjustment (total)	34.03	33.84	0.10
Bell adjustment (social)	12.24	11.40	0.01

TABLE II.—COMPARISON OF TWO-YEAR SUCCESSES WITH FIRST-YEAR FAILURES

Measure	Success mean (n = 117)	Failure mean (n = 47)	D/σ_D
Thurstone	188.35	144.21	7.06
High-school G.P.A	3.04	2.30	11.60
Iowa comprehension	146.19	122.12	6.00
Iowa rate	28.96	25.96	3.15

B. Analysis of Records of Students Finishing One Year.—(1) Correlations with college G.P.A.

The variabilities of the students' scores about the averages of the Thurstone, high-school G.P.A., and especially of the Iowa scores, are very great. Consequently correlations between college G.P.A. and the various measures were run, with results as shown in Table III. The correlation of .707 between high-school G.P.A. and college G.P.A. for a fairly homogeneous group, the graduates of one high school, is strik-

TABLE III.—CORRELATIONS WITH COLLEGE G.P.A.

Measure	r	N	PE_r
High-school G.P.A. (Stockton High graduates only)	.707	125	.030
Thurstone (Stockton High graduates only)	.441	125	.072
Iowa Comprehension (Minus orientals)	.276	241	.040
Iowa rate (Minus orientals)	.148	243	.042

ingly high. The other correlations are similar to those found by other investigators, including the correlation between Thurstone score and college G.P.A. for the graduates of one high school. The Iowa correlations are so low that the test is considered valueless as a predictive aid.

(2) Partial and multiple correlations, Stockton High School Graduates only.

Because of the difference in grading standards of various high schools, the graduates of Stockton High School who completed one year at the Junior College were selected as a group for the study of the relative predictive values of high-school grades and Thurstone scores. The results are presented in Table IV.

TABLE IV.—CORRELATIONS FOR STOCKTON HIGH GRADUATES[1]

(n = 125)

COLLEGE G.P.A. AND	r
High-school G.P.A.	.707
Thurstone	.441
Multiple r Thurstone—High-school G.P.A.	.732
Partial r High-school G.P.A. (Thurstone k)	.648
Partial r Thurstone (High-school G.P.A. k)	.261

[1] Means and standard deviations for these measures are
College G.P.A., M = 1.185, s.d. = .757
Thurstone, M = 169.12, s.d. = 51.0
High-school G.P.A., M = 2.641, s.d. = .617
r Thurstone and High-school G.P.A. = .382
Regression equation is $X_1 = .0029X_2 + .771X_3 - 1.342$
PE est. in multiple R = .348

It is seen that the high-school grades are far superior to Thurstone scores as a single guide to probable college grades, and that the high-school grades taken with the Thurstone scores furnish the best predictive combination.

TABLE V.—CORRELATIONS WITH COLLEGE G.P.A., TWO[1] YEARS

(n = 185)

COLLEGE G.P.A. AND	r
Thurstone	.465
High-school G.P.A.	.668
Multiple r Thurstone—High-school G.P.A.	.704
Partial r High-school G.P.A. (Thurstone k)	.593
Partial r Thurstone (High-school G.P.A. k)	.290

[1] Means and standard deviations for these measures are
College G.P.A., M = 1.299, s.d. = .709
Thurstone, M = 175.60, s.d. = 52.02
High School G.P.A., M = 2.798, s.d. = .371
r Thurstone and High School G.P.A. = .401
PE est. in multiple R = .349

C. Analysis of Students Completing Two Full Years.—Correlations were determined for all students who remained in college for two full years between college G.P.A. and high-school G.P.A., and college G.P.A. and Thurstone score, as shown in Table V. Again we see that the high-school G.P.A. has a higher predictive value than has the Thurstone, even though the high-school grades come from many different schools.

A comparison of the multiple correlations for one and for two years' college work indicates that the first-year grades are predicted more precisely and with a slightly smaller probable error than are the full two years' grades.

VII. CONCLUSIONS, RECOMMENDATIONS, AND SUMMARY

A. Purposes of the Study and the Answers Are: (1) What characteristics differentiate those students who complete the first college year with passing grades from those who do not?

Average high-school grades differentiate most clearly between succeeding and failing students, the American Council on Education test (by Thurstone) differentiates rather well, the Iowa Reading test comprehension and rate scores somewhat, and the Bell Adjustment Inventory scores not at all.

Since students do not drop out of the public junior college because of grades, it being possible for them to remain in school for two years with minus grade point averages, the above question is synonymous with the question, "What measures differentiate students remaining in college two full years from those who drop out before?"

(2) What are the predictive values of the entrance tests?

The high-school G.P.A. is outstanding among the measures available at the time of registration for its predictive value. The correlation between high-school G.P.A. and college grades for one year was found to be .707; that for two years .668. The Thurstone test scores correlate much lower with college G.P.A., the values being .441 and .465, respectively, for one and two years. A combination of the two measures, high-school grades and Thurstone scores, predicts college G.P.A. to a value of .732 for one year and .704 for two. These values are very slightly higher than those for high-school G.P.A. alone. The Iowa Reading Test is practically useless in prediction.

B. Recommendations.—It is interesting that the predictive values for high-school grades and for general ability tests in this junior college are greater than have been found in the majority of similar studies

throughout the country, and that the combination predictive value of high-school grades and general ability test is higher than the combination of any two measures as found by other investigators. These higher values are probably due to three factors; first, that the high-school G.P.A.'s used in the determination of correlations for the first year of college were derived from one high school, and presumably were based on more standard grading conditions than obtained in most studies; secondly, that the junior college has a greater variability of abilities among its students than have regular four-year colleges; and thirdly, that the public junior college guidance program is not one of exclusion of students who fall down in grades. As has been pointed out,[1] the use of tests in personnel research "to identify early those students whose subsequent scholarship will be low, and to divert them into types of academic or non-academic training more in line with their abilities," serves to lower the correlation between those test measures and grades. The youth of the organization of the college in which this study was conducted, and its slight reliance on entrance material in the guidance of students, has served to increase the correlations. As the measures available at the time of registration are more and more used, the correlations will doubtless drop. It is well for the guidance program to have statistical analyses before the program itself takes more definite shape with respect to tests.

The greatest interest of the present study to public junior college counsellors is likely to be in the practical application. The counsellor wants to know what to do with the entrance measures of a student, when the student and the record are before him.

The recommendation to the counsellor, growing from the present research, is the following: Consider first the high-school record of the new student. Disregard details on the nature of his high-school course, giving chief attention to the grade point average. Give second consideration to the Thurstone score, since a student with high ability who has not done well in high school may (although he probably will not) begin to do good work in college.

The counsellor's obligation to the student seems to lie in two fields: First, to direct the student toward those courses for which he is most capable, and, secondly, to prepare the student for his most probable college grades. For example, if the range of probable college grades is about a D, the counsellor should as a rule guide that student away from

[1] (18).

the abstract courses, the college-preparatory courses, and the purely theoretical courses, toward the concrete, terminal, and practical courses. Moreover, the counsellor should start discussing various values with the student so that when the time comes that the student sees his grade card amply filled with D's and F's, with but few C's and no A's, he is not overcome with a sense of failure. Such a student must be led toward objectives other than four-year college, and toward a realization of values other than high college marks. This type of adjustment cannot occur after the student has earned the low grades, but must be built up over the two-year period.

On the other hand, if the student's probable grade point average as predicted is about B, that student should be encouraged to look toward senior college, to value high college marks, and to take theoretical and preparatory courses. This second student may well be encouraged, for example, by his counsellor in borrowing money for continuing in college, while the first student may not.

The recommendation to the Personnel Committee, based on this study, was that the use of the Iowa Reading Tests be discontinued, and that the Coöperative Tests of General Culture be substituted. The reason for the inclusion of the latter is that their scores may have a higher predictive value than do high-school grades. These tests, being definitely achievement tests, measure what the student has learned in high school.

Such a study as this, being a follow-up, must be stopped somewhere, although it is never really finished. Other questions suggest themselves: What happens to students who quit junior college? How do public junior college graduates do in senior college? What best predicts their success or failure in senior college? Do students receiving frequent counselling adjust better in college than do those receiving less counselling? Do students of low (tested) ability achieve better in some courses than in others?

BIBLIOGRAPHY

(1) Douglass, H. R.: "Selecting Good College Risks." *School and Soc.*, Vol. XXXV, 1932, pp. 140–147.

(2) Edds, J. H., and McCall, W. M.: "Predicting the Scholastic Success of College Freshmen." *J. Educ. Res.*, Vol. XXVII, 1933, pp. 127–130.

(3) Eurich, A. C.: "College Failures." *School and Soc.*, Vol. XXXVII, 1933, pp. 692–696.

(4) Ficken, C. E.: "Predicting Achievement in the Liberal Arts College." *School and Soc.*, Vol. XLII, 1935, pp. 518–520.

(5) Fleming, E. G.: "College Achievement, Intelligence, Personality, and Emotion." *J. Appl. Psychol.*, Vol. XVI, 1932, pp. 668–674.
(6) Gladfelter, M. E.: "The Value of the Coöperative English Test in Prediction for Success in College". *School and Soc.*, Vol. XLIV, 1936, pp. 383–384.
(7) Harris, Daniel: "The Relation to College Grades of Some Factors Other Than Intelligence." *Archiv. of Psychol.*, 1931, No. 131.
(8) Johnston, J. B., and Williamson, E. G.: "A Follow-Up Study of Early Scholastic Predictions in the University of Minnesota." *School and Soc.*, Vol. XL, 1934, pp. 730–738.
(9) Landry, H. A.: "The Relative Predictive Value of Certain College Entrance Criteria." *J. Exp. Educ.*, Vol. V, No. 3887, 1937, pp. 256–260.
(10) Nelson, M. J.: "A Study in the Value of Entrance Requirements at Iowa State Teachers College." *School and Soc.*, Vol. XXXVII, 1933, pp. 262–264.
(11) Schmitz, S. B.: "Predicting Success in College: A Study of Various Criteria." *J. Educ. Psychol.*, Vol. XXVIII, 1937, pp. 465–473.
(12) Scott, C. M.: "Background and Personal Data as Factors in the Prediction of Scholastic Success in College." *J. Appl. Psychol.*, Vol. XXII, 1938, pp. 1, 42–49.
(13) Segel, David: *Prediction of Success in College.* U. S. Department of the Interior, Bulletin 1934, No. 15.
(14) Stuit, D. B.: "Differential Characteristics of Superior and Inferior Students." *School and Soc.*, Vol. XLVI, 1937, pp. 733–736.
(15) Wagner, M. E.: *A Survey of the Literature on College Performance Prediction.* Univ. of Buffalo Studies, Vol. IX, 1934, pp. 194–209.
(16) Wagner, M. E.: *Prediction of College Performance.* Univ. of Buffalo Studies, Vol. IX, 1934, pp. 125–144.
(17) Wagner, M. E., and Strabel, E.: "Homogeneous Grouping as a Technique for Improving Prediction Coefficients." *School and Soc.*, Vol. XL, 1934, pp. 887–888.
(18) Williamson, E. G.: "The Decreasing Accuracy of Scholastic Predictions." *J. Educ. Psychol.*, Vol. XXVIII, 1937, pp. 1–16.

XUM

STUDIES IN THE PSYCHOLOGY OF MEMORIZING PIANO MUSIC. VI: A COMPARISON OF TWO FORMS OF MENTAL REHEARSAL AND KEYBOARD OVERLEARNING

GRACE RUBIN-RABSON

New York City

From the beginning of this series of studies in memorizing piano music, the efficacy of analytical study preliminary to keyboard completion has been incorporated into every experimental design. A theoretical distinction is drawn between this analytical study and "mental rehearsal" or "imaginary practice" of the material when it is already learned or well on its way to a given criterion. Practically, however, the two may coalesce when within even a short study period very gifted subjects not only complete the analysis but carry the learning to such a degree of fluency as to require only one or two keyboard trials to meet the standard.

In the experiment directly preceding this one,[12] these study periods were extended to several contrasted lengths in order to encourage the transition from analysis to learning and to evaluate the advantages of further silent learning in insuring economy in keyboard trials and in retention. As indicated by transcriptions of the memorized material, the medium study period produced one hundred per cent learning and the longest period allowed time for a considerable amount of overlearning in addition. Gauged by the keyboard trials, the longer study periods proved economical but not in proportion to the time invested; gauged by amount of retention over a period of two weeks, there were no economies in any one of the three varied lengths.

Perry,[6] observing the relative efficiency of actual and "imaginary" practice, concludes that "imaginary" practice is effective in improving the score in a variety of tasks and that it is more effective in tasks which consist of learning facts that may be observed without actual movement.

In memorizing piano music the preliminary study period emphasizes the intellectual as opposed to the more obviously kinaesthetic aspects of the learning[7] and stresses the values of purposeful as distinguished from incidental learning.[1] It would be an overemphasis, however, not to recognize that even during silent study the tonal images and the organization of spatial units innervate some covert muscle response. That such response is in no way equivalent to actual performance is illustrated by the fact that even when the material can be transcribed

XUM

verbatim from memory a few keyboard trials are required to meet the learning criterion.[12]

THE PROBLEM

Since the extension of the preliminary study period past the analytical stage proved relatively uneconomical, several other time distributions were investigated in the present instance. In each case, five minutes was allowed for analytical study before the keyboard trials.* In the first method, four extra minutes were inserted mid-way during the keyboard learning for mental rehearsal. In the second, the four minutes of mental rehearsal came at the end of the keyboard learning; and in the third, four minutes of extra keyboard trials were added to the finished keyboard learning. Stated briefly, they are:

A. Five minutes of analytical pre-study, five keyboard trials, plus four minutes of mental rehearsal, plus keyboard trials to the criterion of perfect memorized performance.

B. Five minutes of analytical pre-study, keyboard trials to the above criterion, plus four minutes of mental rehearsal.

C. Five minutes of analytical pre-study, keyboard trials to the above criterion, plus four minutes of extra keyboard trials.

THE EXPERIMENT†

Subjects.—The nine adult subjects were all skillful pianists, most of them teachers of piano and possessing at least the minimum amounts of instrumental and theoretical experience required for professional musicians. With the exception of one older subject, the age range was twenty-one to twenty-five years. That the subjects were of unequal ability is witnessed by the average learning and relearning trials which varied from 8.3 to 17, and from 6.3 to 10.5.

Experimental Materials.—Nine short but complete musical periods ranging from five to ten measures in length were adapted from unfamiliar piano materials by Zipoli, Pasquini, Galuppi, Frey, Dohnanyi, Gal and Toch. Extremely varied in style and harmonic range, several of them presented a challenge to even the most gifted subjects.‡

---

* For a rationale of "analytical study" see previous studies [7, 8, 13].

† The experiment was conducted in the studio of the writer in New York City during the Summer of 1940.

‡ It was essential to select materials of sufficient difficulty to guarantee against completion of the learning in the *A* method within the first five trials after five minutes of analytical study.

Experimental Design.—The Latin-Greek square, composed of three smaller squares, served again, as in previous experiments, to control the several rotating variables of subjects, methods, materials, and order. The nine subjects, divided into three groups, each performed the experiment three times, giving a total of eighty-one learnings or twenty-seven for each of three methods.*

Procedure.—The five minutes of analytical study was accomplished at ease away from the keyboard. At the piano no attempt was made to regulate the speed of playing except that extreme deviations and erratic tempi were discouraged.† No devices were employed to encourage speedier learning since these have been proved ineffective.[11]

For the *A* method, after five trials the subject was asked to perform the material mentally with eyes closed, to maintain the image of the notes as firmly as possible, and to refer to the music only when there was confusion or uncertainty in his mental performance. After four minutes of these mental trials, the learning was completed at the keyboard to the criterion of one smooth memorized trial.

In the *B* method, the trials after the five minutes of study were continued to the criterion, after which the already learned material was rehearsed mentally for four minutes.

The *C* procedure was like the *B* until the learning criterion was met, after which the keyboard trials were continued for four minutes.

All the learnings were massed-trial learnings[9] and were accomplished by the "whole"[10] approach. For each subject, the nine learnings were distributed over three successive days. Exactly two weeks later the relearning was done without study periods to measure the relative retention offered by the three learning methods. A second relearning occurred seven months after the learning to estimate the residual benefits.

ANALYSIS OF DATA

For both the learning and relearning trials the critical ratios of the mean differences of the methods are considered a reliable evaluation of any intrinsic advantages. As a further check on their dependability,

---

* For the basic form of this design, and a more thorough rationalization of its use, see [12]. In the present instance the rotation of the compositions was improved.

† The imposition on the pianist of an arbitrary tempo or the necessity of conforming to the clicking of a metronome can prove very confusing and the distraction may far outweigh the value of the extra precision in the laboratory situation.

XUM

the analysis of the variance indicates their significance as well as that of the other variables in the total variation from the mean.*

The learning trials in Table I show a reliable saving for the four minutes of mental rehearsal inserted mid-way. How close was the estimate of five trials as the mid-way point can be seen from the total of 9.89 for this method. For the several subjects this point of learning can not be considered the same, since undoubtedly the better learners were closer to the criterion after five trials than were the less gifted. In support of this assertion, a rough comparison between the most and

TABLE I.—MEANS AND STANDARD DEVIATIONS OF THE LEARNING TRIALS, THE RELEARNING TRIALS AFTER TWO WEEKS AND AFTER SEVEN MONTHS AND THE SIGNIFICANCE OF THE DIFFERENCES OF THE MEANS

Method	Mean	SD	Difference of means	SE difference	Ratio
Learning Trials					
A—Four-minute mental rehearsal mid-way	9.89	3.68	*A-B*, −3.78	1.44	2.63
B—Keyboard learning plus four-minute mental rehearsal	13.67	6.42	*B-C*, 1.93	1.69	1.14
C—Keyboard learning plus four-minute keyboard trials	11.74	5.76			
	10.15	2.89	*A-C*, −1.85	1.33	1.39
Relearning Trials after Two Weeks					
A—Four-minute mental rehearsal mid-way	7.93	2.54	*A-B*, −1.07	.74	1.45
B—Keyboard learning plus four-minute mental rehearsal	9.00	2.81	*B-C*, 1.59	.74	2.15
C—Keyboard learning plus four-minute keyboard trials	7.41	2.60	*A-C*, .52	.71	0.73
Relearning Trials after Seven Months					
A—Four-minute mental rehearsal mid-way	8.11	3.06	*A-B*, − .07	.83	0.08
B—Keyboard learning plus four-minute mental rehearsal	8.18	2.92	*B-C*, .40	.88	0.45
C—Keyboard learning plus four-minute keyboard trials	7.78	3.48	*A-C*, − .33	.82	0.40

* For a more thorough discussion of the use of variance in this connection, see [7].

the least gifted learners proves that the added study period netted the three most gifted a thirty per cent saving over *B*, while for the three least gifted the saving was only fifteen per cent.

The differences between *B* and *C* must be attributed to chance, since by both these methods the procedure to the criterion was identical. In view of this fact, for a more reasonable evaluation of the true situation, the critical ratios of 2.63 between *A* and *B* and of 1.39 between *A* and *C* should be tempered by assuming a theoretical average for *B* and *C*. This would alter the ratio to 2.03 between *A* and *B* and *A* and *C*, which, though not as significant, possibly lies closer to actuality. An average of 10.15 trials was accomplished in the four extra minutes of keyboard trials by the *C* approach, roughly a one hundred per cent overlearning.

The standard deviation of the mid-way study method is not only the smallest, but its proportional decrease is greater than the decrease of the mean.

Undoubtedly the mid-way study method offers a real economy in keyboard trials. Here, as in the case of the experiment described at some length above,[12] this economy must be further evaluated by other criteria, since within the four-minute study period the little compositions could have been repeated approximately ten times.

Table I presents a summary of the amount of retention provided by the three learning methods. The added keyboard trials after bringing the memorizing to the criterion are reliably better than an equivalent amount of mental rehearsal. For skillful musicians such as these subjects, the clarity and precision of the tonal, keyboard, and printed note image after thorough learning may be so vivid as to make the mental rehearsal at least comparable to actual keyboard performance. That it was not as effective for retention is indicated by the ratio of 2.15 for the differences of the means. This is contrary to the finding of Kovacs,[8] who maintained that mental practice after learning was more productive than an equivalent amount of keyboard repetition.

Most interesting is the mean of the *A* method which brought the learning only to the one hundred per cent level and allowed for neither mental nor keyboard overlearning. What the percentage of overlearning might be in the *B* method is impossible to discover under these circumstances but in the *C* method it proved to be about one hundred per cent. This amount of overlearning, however, was no more economical than no overlearning at all, as is shown by a ratio of 0.73 between *A* and *C*.

XUM

The relative advantages of the methods in retention for the three most and the three least gifted were about the same. This conclusion is supported by the narrow range of the standard deviations. An analysis of the experimental materials ranked according to difficulty shows some tendency toward a superiority in the *A* method for the three easiest and the three of medium difficulty, but for the three most difficult, *C* has some extra merit. Of the three most difficult compositions one fell into this category only because of its greater length, while the other two were short but were written in the modern harmonic idiom. The "Klang" of the tone clusters in this idiom is still elusive and does not lend itself to as clear an auditory image as does the color of the old familiar harmonies. Modern harmonies may, therefore, profit to a greater extent from the actual physical spacing on the keyboard.*

TABLE II.—ANALYSIS OF THE VARIANCE OF THE LEARNING AND FIRST RELEARNING TRIALS, SHOWING THE PERCENTAGE OF THE TOTAL DUE TO EACH VARIABLE, AND A COMPARISON OF THE RATIO OF THE VARIABLE AND THE EXPERIMENTAL ERROR WITH FISHER'S ONE PER CENT VALUE

Variable	Fisher's one per cent value	Learning trials		Relearning trials	
		Ratio	Per cent	Ratio	Per cent
Subjects	2.82	7.2	23	6.4	26
Methods	4.98	8.8	28	5.6	23
Compositions	2.82	11.9	38	8.5	35
Order	4.98	2.0	6	3.1	13
Error	....	....	5	...	3

Seven months after the original learning the subjects were recalled for a second relearning. All superiorities in the methods disappeared (Table I). The means approximated each other closely and, on the whole, apparently not much more was lost after seven months than

* As might be expected the more difficult compositions show a greater relative retention than the easier ones. The correlation of rank order for learning and relearning of the materials is 0.87. Measured by speed of learning, the subjects maintain a fairly consistent ranking in learning and relearning, the index being always high, in this case, 0.71.

after two weeks, although the consistently larger standard deviations indicate, presumably, that more was retained by the more able and less by the less able learners.*

The analysis of the variance of the learning and first relearning trials appears in Table II. The three variables of subjects, methods, and compositions play an active and significant part in the total variation. The ratios of each variable and the experimental error are greater in every case than Fisher's one per cent value,[4] that value which might be exceeded in random sampling from a homogeneous population once in a hundred trials. Both the order of learning and the error due to uncontrolled variables are insignificant and only a small part of the whole deviation from the mean. The methods† themselves are responsible for twenty-eight per cent of the learning and twenty-three per cent of the relearning variance and the comparison with Fisher's value lends reasonable support to the significance of the differences of the means.

Only tentative hypotheses can be advanced at this time in explanation of the superiorities of the mid-way study method. The interruption for mental rehearsal may provide a literal "pause that refreshes," due to the cessation of hand movements and an opportunity to think through and reorganize points of confusion without the necessity of maintaining an unbroken rhythm. It may be, too, a kind of "distributed" practice, a procedure previously found advantageous[9] (but exclusively for the least able learners).

For the learner, the four-minute study before the criterion is reached has a different qualitative value from the four minutes of study after it. Mid-way the prestige factors implicit in the situation stimulate his optimum performance. With the learning already accomplished, even with the best will to coöperate, the mental review may be executed with some degree of relaxation.

This may be equally true of the keyboard trials after learning. The fact that neither these nor the extra trials required to reach the criterion exerted any appreciable effect on retention seems to militate against the belief that memorizing piano music is primarily kinaesthetic. When

* The correlation between both relearnings is 0.76.

† The fact that to the subjects and the compositions is ascribable a larger percentage of the total in no way detracts from the significance of the methods. In the future it may be possible on the basis of a more ample selection of already utilized subjects and experimental materials to reduce the contribution of both these variables.

the learning has been completed by the setting together of logical patterns already analyzed and prepared, apparently no particular value lies in repetition past this point. This agrees with Cuff's[2] findings that the "results exhibit no definite relation between additional repetitions and saving" and supports his denial of the assertion[3] that "the frequent passage of a neural impulse somehow leads to a fixation of a habit."

SUMMARY AND CONCLUSIONS

As indicated by previous experiments, an analysis of piano material preliminary to completing memorization at the keyboard proves superior to confining the learning to the keyboard. Extending this analytical period to learning or overlearning offers no economies for retention although it reduces the keyboard trials in learning.

The present experiment maintains the preliminary analytical study, but attempts to evaluate a further period of study or mental rehearsal situated at points in the learning when some keyboard learning has already been accomplished. Toward this end three procedures were compared, all of which incorporated a five-minute period of analysis preliminary to any keyboard trials. The first introduced a four-minute period of mental rehearsal or "imaginary" practice after five keyboard trials or mid-way in the learning, then continued the keyboard trials to reach the criterion of perfect memorized performance. The second carried the keyboard trials to the criterion, then added four minutes of mental rehearsal. The third, like the second, reached the keyboard criterion, then added extra keyboard trials for four minutes.

The mid-way period of mental rehearsal or "imaginary" practice proved reliably superior to the other time distributions. It not only reduced the required keyboard trials, but achieved retention as good as that offered by the four minutes of extra keyboard trials, which, incidentally, resulted in an approximate one hundred per cent overlearning. The four minutes of mental rehearsal placed after the learning is, apparently, an inferior procedure. It not only required more keyboard trials, but produced less retention than either of the other methods. In the second relearning after seven months these differences disappeared. The means, however, were close to the means of the first relearning, but the standard deviations were considerably larger, indicating presumably that little more was lost after seven months than after two weeks and that the less able learners retained less than the more able.

XUM

Several hypotheses may be adduced in explanation of the superiority of the mid-way study method:

First, it is a type of distributed practice, which, though it offers no actual rest, nevertheless provides a period when the cessation of hand movements relieves the necessity for maintaining an unbroken sequence, allows further analysis and reorganization of points of confusion and presents a "re-seeing" of the small musical figures against the general background.

Second, the intensity involved in reaching a learning goal is probably not duplicated in a mental review after the goal has been reached.

Third, when the learning has been constructed on a foundation of logical materials previously well analyzed, a smooth memorized performance is witness to its completion, and no extra drill, either mental or actual, carries it beyond this point.

For practical purposes it might, therefore, be recommended that piano students prepare by thorough analysis an amount of material commensurate with their experience and ability, continue the learning at the keyboard until the hand and tonal patterns have become established, then pause for mental rehearsal or "imaginary" practice until the mental performance can be accomplished smoothly. When, after this, the keyboard practice reaches the criterion of smooth memorized performance, the learning is completed and will probably profit little by continued repetition past this point. Extra trials might be distributed more profitably over subsequent practice periods to restore the learning to the original standard.

BIBLIOGRAPHY

(1) Bromer, J. A. and Meisel, H. S.: "A Comparison of Incidental and Purposeful Memory for Meaning and Nonsense Material." *Psych. Bull.*, Vol. XXXVII, 1940, p. 579.

(2) Cuff, N. B.: *The Relation of Overlearning to Retention.* Geo. Peabody Coll. for Teachers, Contr. to Educ., No. 43, 1927.

(3) Cuff, N. B.: "The Law of Use." *Jour. Ed. Psych.*, Vol. XX, 1929, pp. 438–447.

(4) Fisher, R. A.: *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh, 1934.

(5) Kovacs, S.: "Untersuchungen über das musikalische Gedächtnis." *Zeit. f. ang. Psych.*, Vol. XI, 1916, pp. 113–135.

(6) Perry, H. M.: "Relative Efficiency of Actual and Imaginary Practice in Five Selected Tasks." *Archives of Psychology.* No. 243, 1939.

(7) Rubin-Rabson, G.: "The Influence of Analytical Pre-study in Memorizing Piano Music." *Archives of Psych.*, No. 220, 1937.

XUM

(8) Rubin-Rabson, G.: "Studies in the Psychology of Memorizing Piano Music. I. A Comparison of the Unilateral and the Coordinated Approaches." *J. Educ. Psych.*, Vol. xxx, 1939, pp. 321–345.

(9) Rubin-Rabson, G.: "Studies in the Psychology of Memorizing Piano Music. II. A Comparison of Massed and Distributed Learning." *J. Educ. Psych.*, Vol. xxxi, 1940, pp. 270–284.

(10) Rubin-Rabson, G.: "Studies in the Psychology of Memorizing Piano Music. III. A Comparison of the Whole and the Part Approach." *J. Educ. Psych.*, Vol. xxxi, 1940, pp. 460–476.

(11) Rubin-Rabson, G.: "Studies in the Psychology of Memorizing Piano Music. IV. The Effect of Incentive." *J. Educ. Psych.*, Vol. xxxii, 1941, pp. 45–54.

(12) Rubin-Rabson, G.: "Studies in the Psychology of Memorizing Piano Music. V. A Comparison of Pre-study Periods of Varied Length." *J. Educ. Psych.*, Vol. xxxii, 1941, pp. 101–112.

XUM

SCHOLASTIC STATUS OF ACHIEVERS AND NON-ACHIEVERS OF SUPERIOR INTELLIGENCE

MILTON J. COHLER

Cleveland School, Chicago

That children of high intelligence generally have high scholastic achievement has been well established by numerous studies. However, accomplishment quotients for groups of children of high intelligence are usually less than 100;[1] the average achievement age of a group of bright children is generally less than their average mental age.

From the elementary school through the university, significant discrepancies have been noted between intelligence and scholarship, among individuals of high intelligence. This has been found not only in ordinary schools; but in schools utilizing individual instruction, in special classes for the gifted, and among children with a favored position in society. It was in an attempt to learn more about these disparities between intelligence and achievement that the writer undertook the study, a part of which is presented in this article.

THE PROCEDURE

The subjects for this study were pupils enrolled in twenty-one Chicago schools. In these schools all the children of grades 6*B* to 8*B*, inclusive, who had achieved IQ's of 115 or higher on previous tests of intelligence were given the Otis Self-Administering Test of Mental Ability—Intermediate Examination. Every child whose IQ on the Otis test was less than 120 was then eliminated. The remaining pupils were given the reading comprehension, arithmetic computation, arithmetic reasoning, and language usage tests of the Modern School Achievement Test.

All raw scores on intelligence and achievement tests were then converted to standard scores, utilizing data which the authors of the tests had obtained from a random sample of pupils. To secure a standard achievement score for each pupil, and at the same time avoid weighting arithmetic too heavily, an average for the two arithmetic tests was first computed. Then, this score was used in combination with the anguage and reading scores, respectively.

---

[1] Witty, Paul A., and Jenkins, M. D.: "The Educational Achievement of a Group of Gifted Negro Children." *Journal of Educational Psychology*, Vol. xxv, Nov., 1934, pp. 585–597.

Next, each pupil's Otis standard score was subtracted from his achievement standard score to obtain the standard score difference, which will be referred to hereafter as the "sigma difference." Thus, a large negative sigma difference indicates poor "relative achievement"; a sigma difference with a small negative or zero or positive value, indicates good "relative achievement." The three hundred seventy-five pupils who took both the Otis and achievement tests were then listed in the order of their sigma differences and divided into three groups, approximately equal in number. The one hundred twenty pupils in the middle group were called "intermediates" and were not studied further. The one hundred twenty-seven pupils with the smallest sigma differences were called "achievers," and the one hundred twenty-eight pupils with the largest sigma differences were called "non-achievers."

Achievers and non-achievers were studied further by questionnaires, tests, ratings, and school history. For the purposes of this article, only data from the two tests named above, plus information from teachers' ratings of scholarship and school history, will be given.

THE DATA

I. Intelligence Test Data

The average IQ's of all the pupils to whom both Otis and achievement tests were administered, are given below in Table I. Critical ratios were not calculated because of the anormal character of the distribution. Hence, ranges rather than standard deviations are given.

TABLE I.—AVERAGE IQ'S OF ALL PUPILS FOR WHOM ACHIEVEMENT DATA ARE AVAILABLE

Group	Mean IQ	Range	*N*
Achievers	124.67	120–139	127
Intermediates	125.51	120–142	120
Non-achievers	127.60	120–138	128
Entire group	125.86	120–142	375

It will be noted that the average IQ goes up as relative achievement goes down. Although the table does not classify the pupils according to sex, the same tendency is entirely consistent within each sex group. The average IQ of the non-achievers was three points higher than that

of the achievers. This difference of three IQ points in favor of the non-achievers is statistically significant.[1]

As a further check each one of the three groups was divided into three IQ groups: 120–124, 125–129, 130 and up. By this classification the inverse relationship between relative achievement and IQ shows more clearly. In the lowest IQ category forty-two per cent are achievers and twenty-five per cent non-achievers; in the top category seventeen per cent are achievers and sixty-one per cent non-achievers; in the middle category roughly one-third fall into each of the three achievement groups.

II. Achievement Test Data

Although there was an inverse relationship between relative achievement and IQ, there was a positive relationship between achievement score and mental age score ($r = +.58$). By comparison with a random sample of pupils in the same grades, even the non-achievers (the lowest third of our group) had a high average in academic achievement.[2] An examination of Table II reveals that only in arithmetic were the non-achievers slightly below the average of the older group found in a random sample.

The averages of the achievers were significantly higher than those of the non-achievers in each subject as well as on the whole battery. Thus the high critical ratios shown in Table II were calculated from data which are internally consistent. Since there was no difference in chronological age between the groups, and the mean IQ of the non-achievers was slightly higher than that of the achievers, the difference in relative achievement is even greater than the standard score difference shown in Table II.

It will be noted that the difference in favor of the achievers is greatest in arithmetic and least in reading comprehension. A study of academic interests showed that more achievers than non-achievers

---

[1] There was no difference between the groups in chronological age

$$\left(\frac{\text{Diff.}}{\sigma\ \text{Diff.}} = 0.04\right).$$

Therefore, Otis standard scores were substituted for IQ's. When this was done the $\frac{\text{Diff.}}{\sigma\ \text{Diff.}}$ was 4.40.

[2] A standard score of 50 represents average attainment.

XUM

listed arithmetic among their best liked subjects. There was no significant difference between the groups in their preference for reading, or in the amount of time spent in this pursuit. In each group ninety-five per cent of the children said they liked to read.

TABLE II.—STANDARD SCORES FOR ACHIEVERS AND NON-ACHIEVERS ON THE MODERN SCHOOL ACHIEVEMENT TEST

Subject	Group	Mean	SD	*N*	Mean difference	Diff. / σ Diff.
Reading comprehension	Achiever	66.48	4.62	127	4.07	6.41
	Non-achiever	62.41	5.55	128		
Arithmetic comprehension	Achiever	54.08	6.87	127	9.11	8.63
	Non-achiever	44.97	9.75	128		
Arithmetic reasoning	Achiever	56.28	7.02	127	7.46	7.93
	Non-achiever	48.62	8.04	128		
Language usage...	Achiever	60.67	5.97	127	7.00	9.86
	Non-achiever	54.67	5.35	128		
General achievement	Achiever	60.54	3.89	127	6.54	12.60
	Non-achiever	54.00	4.37	128		

III. Teacher Judgment of Academic Achievement

No detailed estimate of academic attainment was obtained from the teachers, because teachers are too likely to rate pupils in terms of class norms. The writer wanted a rating of achievement related to each pupil's mental ability. Indirectly, however, teachers gave an estimate of general academic attainment in relation to mental ability by answering "yes" or "no" to the question, "Is this pupil generally superior in academic school subjects?" Explanatory notes were added to the question to minimize the halo effect of personality traits and to exclude motor skills from the judgment. Since a child of superior intelligence should be generally superior in academic achievement in order to be an achiever, the answer to this question gave, roughly, the teacher judgment desired.

Teachers rated seventy-three per cent of achievers "yes" and twenty-seven per cent "no." Thirty-five pupils of superior intelligence who achieved well in relation to their intelligence were not rated generally superior in academic work. One of these children was among the eleven best achievers in the whole group. Of the non-achievers sixty-four per cent were rated "yes." Five of the eleven worst non-achievers were included in this sixty-four per cent.

IV. Acceleration and School History

At the time that most of the data for this study were secured the mean grade level for the achievers was 7.57, and for the non-achievers 7.29, a difference of about six-tenths of a semester. Since there was no significant age difference between the groups, we could anticipate that the achievers' mean acceleration would be about six-tenths of a semester greater than that of the non-achievers, if the data were internally consistent. The actual difference found was 0.667 semesters. This difference was statistically significant $\left(\frac{\text{Diff.}}{\sigma\ \text{Diff.}} = 6.40\right)$. The achievers' mean acceleration was 1.40 semesters and that of the non-achievers, 0.73.

A further check was made on the internal consistency of these data. If a real relationship existed between relative achievement and amount of acceleration, this relationship would be apparent within each group, as well as in a comparison of groups. Hence the two groups were combined in a correlation analysis. With chronological age held constant by the method of partial correlation, the Pearson coefficient of correlation between standard achievement score and semesters of acceleration was found to be +.40. In order to be certain that this coefficient was not influenced by IQ, a correlation coefficient was computed between IQ and acceleration, by correlating mental age score with acceleration, while chronological age was held constant. There was no correlation between IQ and semesters of acceleration. Although the data indicate a relationship between relative achievement and acceleration there was one achiever retarded chronologically by one semester and eighteen others who were not accelerated. Twenty non-achievers were accelerated two semesters or more.

Kindergarten attendance was found to be unrelated to relative achievement—about three-fourths of each group had attended kindergarten. There was also no noteworthy difference between the groups in number of schools attended since entering first grade. Whatever may have been lost in transferring, insofar as scholastic achievement is concerned, appears to have been compensated for by the varied stimulation of different schools and teachers.

SOME OBSERVATIONS AND CONCLUSIONS

(1) In summarizing ten studies of bright children, Witty and Jenkins found that the average achievement quotients of nine of the ten

groups was less than 100 $\left(\text{Achievement Quotient} = \frac{\text{Educational Age}}{\text{Mental Age}}\right)$.[1] Corroborating Witty and Jenkins' summary, this study also found disparities between achievement and intelligence which were numerous and often very large. These disparities suggest that achievement is the resultant of numerous forces other than those measured by tests of intelligence, and that levels of expectancy based largely on mental ages are of little value for the bright child.

(2) The steady increase of the disparity between intelligence and relative achievement, with an increase in the mean IQ of the group, indicates that even the limited objectives exemplified by standard academic achievement tests are not being adequately realized in the case of the bright pupil. One does not have to look far for reasons why relative achievement goes down as intelligence goes up. The child of high intelligence has had fewer years of school and life experience than the child of the same mental age with a lower IQ. School instruction, as well as social demands and opportunities, come nearer to exploiting one's fullest capacities, the closer those capacities approximate the mean. Thus the child of high intelligence, favored though he is, has certain handicaps to good relative achievement which must be compensated for by the force of increased drive. Strong motivation, or drive, is influenced by the total educational milieu, of which the school is only a part. The school can best do its part by providing an educational program which will tend to encourage drive. Such a program would emphasize the setting of individual goals, and each child would share actively in the creation of those goals.

(3) One solution that has been often proposed to remedy the disparity between intelligence and achievement among bright children is the use of acceleration to provide more stimulation. It has been suggested by some that acceleration is enrichment, and by others that it is a valid substitute for enrichment, pending the time when educational procedures become sufficiently advanced to provide real enrichment.

In this study, some relationship was found between acceleration and relative achievement. Though some teachers would be inclined to assert that the acceleration was due to excellent achievement, those who have had broad contacts with school administration know that school policies with respect to acceleration vary considerably. School

[1] Witty, Paul A. and Jenkins, M. D.: *Op. cit.*

principals vary in their ability to resist parental pressure for unwarranted acceleration; ambitious parents send their children to summer school to accelerate them; and different teachers' estimates of a given child often disagree. If excellent achievement were the main cause of acceleration, we would not find twenty of our non-achievers accelerated two or more semesters, while nineteen of our achievers were not accelerated at all.

Furthermore, there is a marked relationship between IQ and achievement. We can assume that good intelligence causes good achievement. If good achievement, which was caused, at least in part, by good intelligence, were the cause of acceleration, we would find that IQ and acceleration were related. However, there was no correlation between IQ and acceleration within the IQ range studied. Thus we may conclude that good achievement does not cause acceleration. Another hypothesis may then be advanced. In view of the reasoning given above in (2), it is suggested that a reasonable hypothesis is that acceleration was one of the causes of excellent achievement.[1] Hence, a more general use of acceleration for pupils of superior intelligence may be desirable. The writer would urge caution and limits upon this practice, and that it be followed only after a careful study of each individual from the standpoint of scholarship, personality, and physical characteristics.

(4) Under the conditions of this study it was entirely reasonable for teachers to rate many non-achievers generally superior in academic work; but when we find that thirty-five achievers of superior intelligence were not rated academically superior, it must be concluded that there was a sharp disagreement between standard test results and teacher judgment. The remedy is not to substitute test results for teacher judgment or *vice versa*. It lies rather in training teachers to utilize and interpret standard tests in forming their judgment of children's attainment. It is true that personality, home background, and a host of other factors affect the pupil's academic achievement. Nevertheless, teachers should be trained to make a systematic effort to avoid being influenced by these factors in their judgment of the realized academic performance of the pupil. The child with a personality handicap should not be doubly penalized for it: once in the quality of his achievement, and once again in the teacher's judgment of what he has actually achieved.

[1] See Terman, Lewis M.: *Genetic Studies of Genius*, Vol. I, p. 629.

XUM

(5) The fact that kindergarten attendance and frequency of transfer from one school to another were unrelated to relative achievement, suggests two hypotheses to the writer. Perhaps frequent transfer does not affect bright children because they are very adaptable and quickly see through the slight changes encountered upon entering a new school. Since the educational effort required of them is pitched at a relatively low level in almost all schools, even a change in school would involve no great challenge to their abilities. Studies which would compare bright with average children for the influence of school transfer on relative achievement could settle this point. It would be necessary in such studies to hold socio-economic level and other important aspects of family status constant, since these factors influence both school transfer and relative achievement.

If it were found that even among average children there was no relationship between relative achievement and frequency of transfer, it could be inferred that education lacks a continuity integrated with the normal development of children. If education consisted of vital, meaningful experiences, all woven into a plan to fit the growth and development of children, a school transfer would constitute a serious interruption—a change of motif which could not be dismissed lightly. The writer is not proposing a new difficulty for children. It is merely suggested that school experiences be made so significant that they would become a more vital factor in child development.

XUM

STUDY HABITS ASSOCIATED WITH HIGH AND LOW SCHOLARSHIP

C. GILBERT WRENN AND WILBUR J. HUMBER

University of Minnesota

The measurement of the factors influencing student success and failure has intrigued educational research workers for many years. The use of scholastic aptitude tests to compensate for the unreliability of previous grades in predicting scholarship has been a step forward. It is true that previous grades are still the most efficient single predictor of potential scholarship, but adding scholastic aptitude test scores markedly improves the regression equation. The non-scholastic factors which influence scholarship have not, however, yielded so readily to quantitative treatment. In particular, the study habits of students have been analyzed repeatedly but without great success. One fault of many investigations that have been made in this area is failure to control the related factor of intelligence, so that the results too frequently have been a composite of unknown determination.

Perhaps even more than the research worker, the school counselor or clinical worker has wished to secure some measure of the work habits used by students. With many students the clinical worker is at a loss to know the reason for the discrepancy between measured ability and actual accomplishment. Study habits are suspected to be an important factor, but just what habits are the possible worst offenders is not known. With still other students who know that they need help in their habits of work the problem of knowing where to begin is a real one. It would be best actually to observe the students at work, but the time involved makes this impossible in any but isolated situations. Furthermore, the relation of this factor of study habits to other possible causes of low achievement needs to be known. Various remaining factors that might enter into the failure of aptitude tests and previous grades to account completely for present academic success and failure include differences in scholastic aptitude, length of college experience, number of credits carried, pattern of studies, and sex differences as well as the study habits practiced. It is probably unnecessary to give equal consideration to the separate factors of place of residence and amount of outside remunerative work. One of the most recent and carefully controlled investigations on scholar-

XUM

ship and place of residence[1] reveals that scholarship, when high-school rank is held constant, is not significantly associated with place of residence. Other controlled studies[2] show that except for excessive amounts of work there is little association between the extent of such work and scholarship, when ability is held constant.

In the investigation reported here the writers were in general concerned with the isolation of the factor of study habits and its analysis with reference to high and low scholarship. It was specifically proposed, first, to attempt the validation on a new population of the items used in the Wrenn *Study Habits Inventory*.[3] Secondly, the authors were interested in determining the extent to which study habit items were valid for women as well as men since the *Inventory* was originally standardized on men only. In order to do this, the junior author, using the methods developed in the Stanford University study in 1934–1935, made an analysis of the study habits associated with the "successful" and "unsuccessful" academic careers of University of Minnesota men and women students. This was achieved by isolating the factor of study habits for investigation by holding constant certain other facts which might also contribute to scholarship.

Reports on study habits and attitudes were received from a sample of fifty pairs (twenty-four pairs of women and twenty-six pairs of men) of 1939–1940 Science, Literature and Arts sophomores at the University of Minnesota. These pairs were selected out of a total of one thousand five hundred thirty-three records examined, representing the total freshman Science, Literature and Arts registration for the Fall quarter of 1938–1939. One member of each pair was in the lower forty per cent of his class in honor point ratio and so was considered to be "unsuccessful," while the other was in the upper twenty per cent of his class and was considered "successful." These students were paired on the following variables: (*a*) within one probable error of an individual score on the 1937 American Council on

---

[1] Van Alstine, Frank I.: *The Relation Between the Housing of Students and Their Scholarship at the University of Minnesota.* Unpublished Ph.D. thesis, University of Minnesota Library, 1941.

[2] Umstattd, J. G.: *Student Self Government at the University of Minnesota.* Minneapolis, Minn.: University of Minnesota Press, 1932.

Reeder, C. W. and Newman, S. C.: "The Relationship of Employment to Scholarship." *Educational Research Bulletin*, Vol. XVIII, No. 8, November 15, 1939, pp. 203–214.

[3] Wrenn, C. Gilbert, assisted by R. B. McKeown: *Study Habits Inventory.* Stanford University, California: Stanford University Press, 1934.

Education *Psychological Examination*, (*b*) sex, (*c*) the same major pattern of course work, (*d*) within two credits of the same academic load, (*e*) within one quarter of the same length of academic experience, (*f*) no systematic training in study habits.[1] Having controlled all those factors which have obvious relationships to scholarship except that of the study habits practiced, the investigators concluded that they were in a position to study the uncontrolled variable of the study habits that are associated with scholastic achievement.

The one hundred students whose responses are the basis of this investigation represent an eighty-nine per cent return of the questionnaires which were sent out. These students checked an inventory of seventy-one habits and attitudes of study compiled from various sources, twenty-five of them from the published *Study Habits Inventory*. The directions at the head of the list of items asked the student to put a check mark in the appropriate column after each item which most nearly described whether he "rarely," "sometimes," or "very frequently" practiced that study habit. He was asked to indicate what he did, not what he thought he should do. These responses were anonymous in order to provide conditions that would make it unnecessary for a student to respond in a "socially approved" manner. The student was told on the other hand that the investigator could identify his responses by a code number known only to the research worker. These inventories were sent out and returned within the week of May 4, 1940.

The results were tabulated by determining the percentage of students from the high and low scholarship groups, respectively, who responded to each item in the check list. An item was considered to discriminate between high scholarship and low scholarship students with statistical significance if in any one of the three columns of response, ("rarely," "sometimes," and "frequently") the difference in the percentage of responses between the two groups of students was at least twice the standard error of the difference.[2]

---

[1] The distribution of the A.C.E. scores of the one hundred students did not differ from a normal distribution at the one per cent level of statistical significance. Consequently it is possible, so far as aptitude is concerned, to generalize from this population to similar populations whose ability scores are normally distributed.

[2] The use of Student's "*t*" which would have taken into account the degrees of freedom would have given a more rigorous measure of these differences. The use of such a measure would have tended to increase rather than reduce the number of significant differences observed.

XUM

TABLE I.—STUDY HABITS AND ATTITUDES UPON WHICH HIGH AND LOW SCHOLARSHIP MEN DIFFERED SIGNIFICANTLY

1. I miss important points in the lecture while copying down notes on something which has gone before.
2. I try to summarize, classify, and systematize the facts learned.
3. I try to do some "over-learning," and work beyond the point of immediate recall.
4. I go back and recite the material studied to myself, re-checking any points I find doubtful.
5. I study more with others than by myself.
6. I plan out in my mind the answer to a subjective question (essay type) on an exam before starting to write out the answer.
7. I finish my exams and turn them in before the time is called on the exam.
8. I am among the last to finish and turn in an exam.
9. I get "fussed" and nervous on exams—blow up and can't do myself justice or tell what I know.
10. I skip charts, graphs, and tables in the body of the reading material because they take too much time to understand.
11. I have trouble picking out the important points in the material read or studied; take down material which later turns out to be merely explanatory or irrelevant.
12. I have to re-read material several times—the words don't have much meaning the first time I go over them.
13. I pronounce the words to myself as I read.
14. I make it a habit to look up the meaning of words and technical terms I run across and of whose meaning I am not certain.
15. My periods of study are interrupted by outside interferences such as telephone calls, visitors and distracting noises.
16. I do my studying in the library.
17. I find myself too tired, sleepy, and listless to study efficiently.
18. I find it hard to force myself to finish work by a certain time, or under pressure; work is unfinished, inferior, or not in on time.
19. I find it hard to keep my mind on what I am studying—don't know what I have been reading when I get through.
20. I "rationalize" and find excuses why I cannot study—"have a headache," "think I'm taking the flu," etc.
21. It takes me some time to get settled and "warmed up" to the task of study.
22. My study periods are often too short for me to get warmed up and concentrated.
23. I "take it easy" in the first few days of each new quarter.
24. My time is not wisely distributed; I spend too much time on some things and not enough on others.
25. Too much social life interferes with my school success—dances, dates, and trips.
26. I spend so much time on extra-curricular activities—debating, dramatics, school papers or magazines, dorm, club and fraternity activities, etc.—that my school work suffers.
27. I spend too much time reading fiction, going to shows or enjoying social life for the good of my school work.

XUM

TABLE II.—STUDY HABITS AND ATTITUDES UPON WHICH HIGH AND LOW SCHOLARSHIP WOMEN DIFFERED SIGNIFICANTLY

1. I use a typewriter in writing reports to be turned in.
2. I review everything I have studied regardless of its possible importance.
3. I finish my exams and turn them in before the time is called on the exam.
4. I am among the last to finish and turn in an examination.
5. I get "fussed" and nervous on exams—blow up and can't do myself justice or tell what I know.
6. I pronounce the words to myself as I read.
7. I can't work well where I cannot smoke—have to go out and "have a smoke" when working at the library, etc.
8. I find it hard to force myself to finish work by a certain time, or under pressures; work is unfinished, inferior, or not in on time.
9. It takes me some time to get settled and "warmed up" to the task of study.
10. I try to study too late at night; rob sleep time and meal time for study.
11. I spend too much time reading fiction, going to shows or enjoying social life for the good of my school work.

Table I lists the twenty-seven items upon which the two groups of men differed significantly, and in Table II appear the eleven statistically significant items for the women. It will be noted that the largest share of these items was concerned with examinations, distribution of time, sociability, and concentration, while habits in the areas of study environment and health contributed but few significant items.[1]

Twenty of the twenty-five items selected from the *Study Habits Inventory* proved to be as significant for men at Minnesota as for the men in the earlier investigation at Stanford. All but one of the items (insufficient exercise) was in the same direction of significance as in the earlier publication. Eleven of the twenty-five items from the *Inventory* were significant for the women. Since the latter instrument was standardized on men, this limitation of results for the women can be understood. It does, however, suggest the hazard of using all of the items in the original (or unrevised) *Study Habits Inventory* for diagnosing the study techniques of women students. These sex difference results have been incorporated in the scoring key of the Revised *Study Habits Inventory*.[2]

[1] The amount and direction of the percentage differences between high and low scholarship students in each of the three columns may be ascertained directly from the authors or found, transmuted into scoring weights, in the scoring key of the Revised *Inventory*.

[2] Wrenn, C. Gilbert, assisted by Wilbur J. Humber and R. B. McKeown: *Study Habits Inventory* (Revised). Stanford University, California: Stanford University Press, 1941.

XUM

To date there has been little empirical evidence on the study habits which differentiate men and women. In this study the women contributed less than one half as many significant items as did the men. Only one-fourth of the items significant for men were also significant for women. It is of interest to note that four items were significant for women which did not discriminate between men. From these it is learned that successful women more than successful men use a typewriter for written reports, participate in class discussions, and rob sleep time for study. Unsuccessful women more than unsuccessful men can't study unless they "go out and smoke occasionally."

The inability of the investigation to reveal as many habits differentiating high-scholarship women from low-scholarship women as was true for men suggests that insofar as these items are concerned, women are less variable in study skills than are men. If there are persistent sex differentials in study habits it will be necessary for study habits courses and educational counselors to differentiate their teachings and recommendations. However, such a conclusion is premature until further study confirms or modifies the findings of this investigation.

The writers are aware of the limitations of a procedure which depends upon the subjective judgment of students. However, since direct observation of study habits for any considerable number of students is impractical, the use of such an instrument as the *Inventory* enables the personnel worker or teacher to identify quickly those students whose study habits deviate seriously from the practices of successful students. Having identified these, the counselor is in a position to examine the individual responses of the student and in subsequent interviews to attempt to determine whether such study practices are incidental to or the cause of poor scholarship.

XUM

IQ CORRELATED WITH EQ

HENRY FEINBERG*

Psychologist, Jewish Social Service Bureau, Detroit; Director, Educational and Vocational Clinic, Great Lakes College; Instructor, Psychology, Great Lakes College, Ferndale, Michigan

The purpose of this study is to determine the relationship of the intelligence quotient, as described by such intelligence tests as the Stanford Revision of the Binet-Simon Tests,[7] the Terman-Merrill Revision of the Stanford-Binet Scale,[8] and the Wechsler-Bellevue Intelligence Tests,[9] to the educational quotient as described by the New Stanford Achievement Tests, Forms *V*, *W*, *X*, *Y*, and *Z*.[3]

Other studies have shown positive correlations between the two types of tests, despite the fact that the philosophy underlying the intelligence test differs materially from that underlying the achievement test. The intelligence test, for instance, describes intelligence, "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment" as Wechsler defines it;[9] while achievement attempts to describe how well our subject has dealt with the educational part of the environment. One describes the limits within which one may be expected to deal with an educational problem; the other how well he has accomplished the task.

One may, therefore, expect positive correlations and, perhaps, good correlations between these two types of tests when one is dealing with groups of persons. Gilbert S. MacVaugh stated that "an occasion for an increase in psychometric program efficiency arose in the psychological clinic at the U. S. Penitentiary, Leavenworth, Kansas, in which it was desired to convert the scores made on the intelligence test into educational ages." In examining two thousand male inmates MacVaugh found that the Revised Alpha Examination, Form 5, and the New Stanford Achievement Test, Form *W*, correlated with each other .87, the PE being $\pm$.004. There were no statistically significant differ-

---

* Four students in a course in Interpretation of Psychological Examinations, given at Great Lakes College, Ferndale, Michigan, collaborated in this study, under the direction of their instructor, Henry Feinberg: Mrs. Irene Dugan, Principal, Erin School, East Detroit, Michigan; Mrs. Helen Mae Krause, Teacher, Eastwood School, Wayne County, Michigan; Miss Esther D. Schott, Teacher, Erin School, East Detroit, Michigan; and Mrs. Beth Sheridan, Acting Principal, Eastwood School, Wayne County, Michigan.

XUM

ences found between the means obtained on both the intelligence and the educational test scores when they were compared with a previous survey of the group on the two tests.[5] Henriette Woolf and Christine Lind reported that the New Stanford Achievement Test and the Stanford Binet Scale correlated .80, the PE being .02.[10] St. John made an extensive study on this relationship. He found correlations ranging from very insignificant correlations to positive correlations of a marked degree, with the mean correlation being .56.[6] Harold A. Kohn reported the correlation between the Ohio State Psychological Examinations and the Sones Harry High School Achievement Test to be .82.[4] Arthur Gates divided his subjects into grade categories. The correlations were consistently positive.[2]

There has been a tendency upon the part of some of the local rural school systems to substitute the educational quotient for the intelligence quotient, the practice being based on the belief that, by and large, there was no material difference between the two in describing the ability of the person tested, and that the educational quotient had the added virtue of giving the school staff an indication of the proper grade allocation for the child. Indeed, some of the psychological literature encouraged such thinking. With this problem in mind, we proceeded to discover what our clinical population revealed when their IQ was compared with their EQ.

Seven hundred and fifteen children were examined at the Mental Hygiene Clinic of the Jewish Social Service Bureau of Detroit, Michigan. These children came from children's homes of various religious denominations, from foster homes, from normal homes and from institutions for predelinquents. They had attained 3-*A* grade and had not passed beyond the tenth grade. Their IQ's ranged from 44 to 192 (the distribution will be seen in Table I). As we wish to discover whether we could substitute EQ for IQ in individual instances, we divided the population studied into diagnostic groupings, in which those having IQ's below 70 were classified as mental defectives; those between 70 and 79 as having borderline intelligence; those between 80 and 89, dull or low normal intelligence; 90 to 109, normal intelligence; 110 to 119, bright normal; and those above 120 as being superior children.

If it is possible to substitute one test score for another, such as MacVaugh suggested and as appears to be the practice in some school systems, then one must presuppose that an educational quotient suggestive of a diagnostic level of general intelligence can be adequately

substituted for the intelligence quotient, the instrument used at present to interpret such a level. An IQ of 80, for an example, is ordinarily interpreted as low normal intelligence. Should a clinical subject obtain an educational quotient of 80, the necessity for giving him an intelligence test would, therefore, be ruled out as mere repetition and a waste of clinical time. The data on hand, however, indicate that such a substitution cannot be made.

TABLE I.—CORRELATIONS BETWEEN IQ AND EQ OF CHILDREN ACCORDING TO THEIR DIAGNOSTIC GROUPINGS*

IQ range	EQ range	Number	Correlation	PE ±	SD IQ	SD EQ	Mean IQ	Mean EQ
44– 69	56– 84	53	−.37	.09	4.4	6.3	64	71
70– 79	59– 98	111	−.36	.06	3.6	7.6	75.6	76.5
80– 89	59–102	138	−.11	.05	2.6	7.6	84.6	82.
90–109	58–120	300	.51	.03	5.3	10.4	98	92
110–119	86–123	68	.15	.08	2.0	8.9	113	106
120–102	89–149	45	.76	.04	12.1	11.5	132	118

* The formula used was:

$$r = \frac{N\Sigma xy - \Sigma x \Sigma y}{\sqrt{N\Sigma x^2 - (\Sigma x)^2}\sqrt{N\Sigma y^2 - (\Sigma y)^2}}$$

The advisability of giving the New Stanford Achievement Test to persons having IQ's below 70 may be raised. But the fact that the tests are administered to these persons in the school systems and judgment passed by the school authorities on the ability of these persons on the basis of the results obtained on these tests, appears sufficient reason to include them into our group under study. Of the fifty-three having IQ's below 70, thirty-four had EQ's of 70 and above; and four, 80 and above. In other words, the general impression was that these children either had borderline general intelligence or were dull or low normal in intelligence. Of these one hundred eleven children, IQ 70–79 inclusively, twenty-one had EQ's below 70 and forty above 79. Of the one hundred thirty-eight children, IQ 80–89 inclusively, fifty-nine had EQ's below 80 (forty-nine between 70–79 inclusively, nine 60–69 inclusively, and one below 60); and twenty-five above 89, (twenty-four 90–99 inclusively, and one, 100–109 inclusively). Of the three hundred children having normal general intelligence, (IQ 90–109, inclusively), one hundred twenty-nine had EQ's

XUM

below 90 (one hundred six, 80–89, inclusively; twenty 70–79, inclusively; two, 60–69, inclusively; and one below 60), while fifteen children had EQ's 110–120 inclusively. One hundred forty-four, or almost half, of the children having normal intelligence as described by the standardized intelligence test had EQ's which, if substituted for IQ, would lend themselves to a wrong interpretation of the child's ability. Forty-two of the sixty-eight children whose IQ's were 110–119, inclusively, had EQ's 90–109 inclusively, and three, 80–89 inclusively, while four had EQ's 120–123 inclusively. Twenty-seven of the forty children, IQ's 120 and over, had EQ's below 120, (seventeen, 110–119, inclusively, eight, 100–109, inclusively, one, 90–99 and one 89).

Negative correlations were obtained among the mental defectives, those having borderline general intelligence and among those having low normal general intelligences. There appears to be a negligible positive correlation among the bright normal group while those having normal general intelligence have a marked correlation and the superior children a high correlation between the intelligence test scores and the educational test scores.

It is interesting to note that as the IQ rose there seemed to be a general tendency for the correlation to rise. The one exception to this was in the group 110–120. This phenomenon is not new. Henry Feinberg found a similar tendency when he correlated the Stanford Revision of the Binet Simon Test with the Pinter Paterson Short Performance Scale.[1] Harold St. John seemed to find it in his study of achievement in its relation to other factors.[6]

TABLE II.—CORRELATION BETWEEN MALES AND FEMALES AND THE WHOLE GROUP STUDIED IN THEIR PERFORMANCE ON TESTS OF GENERAL INTELLIGENCE AND ACHIEVEMENT

Item	Range of IQ	Range of EQ	Number	Correlation	PE ±	SD IQ	SD EQ	Mean IQ	Mean EQ
Females.....	65–147	65–131	98	.51	.05	14.96	13.27	101.8	99.7
Males.....	44–192	56–149	617	.82	.008	17.90	14.89	92	87.3
Total.....	44–192	56–149	715	.82	.008	17.55	14.88	93	89

As seen by Table II, in which the correlations between the results obtained on the intelligence test and educational test among the males, females and the entire group of seven hundred fifteen persons are described, the results do not differ from those obtained by Mac-

Vaugh,[5] Henriette Woolf and Christine Lind,[10] Kohn[4] and Arthur Gates.[2] All correlations are highly positive.

Significant differences occurred when we compared females and males in their performance on the intelligence tests and found that the difference of their means was six times the standard error of difference; that on the educational tests the standard error of the differences was three and six-tenths times the difference of their means.

The females in this group revealed no significant difference between their performances on the intelligence test and on the educational test, the ratio being 1.5. The males, however, revealed that the difference of their mean performance on the two tests was 11.46 times the standard error of the difference. The whole group of seven hundred and fifteen persons performed so that difference of their mean performances was more than eleven and eight-tenths times the standard error of the difference.

Three hundred and ninety-six of the above seven hundred and fifteen children would fall into a different intelligence-diagnostic category if we used the EQ to express their general ability.

The discrepancy between the IQ and the EQ of individuals may involve several etiological factors, all of which can not be stated here. Among these are status of the curricula of the school system itself, the emotional condition of the person tested, the interest of the person tested, and the like. For those persons having higher IQ's it is probable that many of the fundamentals in any branch of education, particularly grammar and arithmetic, have not been communicated to them. Among these persons are students who have particular and temporary interests which exclude other interests, so that a person interested in geography may become so absorbed in it that he may neglect to study hygiene, arithmetic and the like. Their scores in these latter subjects would weigh down the score representing the total result. On the other hand the older child with a low IQ may have been placed in a position through the process of promoting because of his age whereby he can accumulate information which he may never have the need to use and which for the most part he will eventually forget. One is reminded of a girl of sixteen who graduated from a high school academic course, and whose IQ was 79. She applied for work as a stenographer, a typist, then as a file clerk. She failed to hold any of these jobs and finally found herself without work. She did not know anything of domestic service in which, she stated, she would be happy. On the achievement test she had a normal score for one who

XUM

is in the third year of high school. But she was rapidly forgetting a great deal of what she had learned "because I never use the stuff." The fact that one acquires information can lead to no prediction that the information is of use in any adequate sense or that it will be retained. The relationship of the IQ to EQ is one in which it is possible to note if one is performing educationally on the level, below the level, or above the level of his general ability. A discrepancy between the performances on the part of an individual should be studied from the standpoint of the causal factors involved.

CONCLUSIONS

(1) The correlation obtained on the whole of the seven hundred and fifteen persons in the group studied above, in their performances on intelligence and educational tests, do not differ from correlations obtained in groups described by other persons.

(2) However, when it is broken down in diagnostic groups there appears to be a negative correlation among the mental defectives, persons of borderline intelligence and persons of dull normal intelligence, a slight correlation among persons having high normal general intelligence, and a substantial correlation among the persons having normal and superior general intelligence.

(3) Significant differences were found between the mean performances of males and females on both the intelligence and educational tests, in the male population between their mean performances on the intelligence and educational tests as well as the mean performances on the two tests among the total population tested.

(4) There appears to be a perceptible rise in correlation with the rise of IQ, although the high normals are an exception to this.

(5) From the evidence we have on hand the EQ cannot be substituted for the IQ.

BIBLIOGRAPHY

1. Feinberg, Henry: "The Stanford Revision of the Binet Simon Tests compared with the Pintner-Paterson Short Performances Scale." *Pedogogical Seminary and Journal of Genetic Psychology*, Vol. XL, March, 1932, p. 499.
2. Gates, Arthur I.: "The Correlations of Achievement in School Subjects with Intelligence Tests and Other Variables." *Journ. of Ed. Psych.*, Vol. XIII, pp. 129–139, 223–235, 277–285.
3. Kelley, Truman L., Giles, M. Ruth, and Terman, Lewis M.: *New Stanford Achievement Test Forms V, W, X, Y, Z.* World Book Company, New York, 1929–1931.

XUM

4. Kohn, Harold A.: "Achievement and Intelligence Examination Correlated with Each Other and with Teacher's Rankings." *Journ. of Genetic Psych.*, Vol. LII, June, 1938, pp. 433–437.
5. MacVaugh, Gilbert S.: "Regression Formula to Predict Educational Ages from Revised Alpha Scores." *Journ. of Applied Psychology*, Vol. XXII, June, 1938, pp. 275–281.
6. St. John, Charles W.: *Educational Achievement in Relation to Intelligence.* Harvard University Studies Vol. 15, Harvard University Press, Cambridge, 1930.
7. Terman, Lewis M.: *The Measurement of Intelligence.* Houghton Mifflin Co., Boston, 1916.
8. Terman, Lewis M., and Merrill, Maud A.: *Measuring Intelligence.* Houghton Mifflin Co., Boston, 1937.
9. Wechsler, David, *The Measurement of Adult Intelligence.* The William and Wilkins Company, Baltimore, 1939.
10. Woolf, Henriette and Lind, Christine: "A Study of Some Practical Considerations Involved in the Use of Two Educational Test Batteries." *Journ. of Ed. Psych.* Vol. XVI, Nov., 1938, pp. 629–634.

XUM

READING GAINS IN A FRESHMAN REMEDIAL PROGRAM AT SYRACUSE UNIVERSITY

DOUGLAS F. PARRY

Department of Psychology, Syracuse University

Syracuse University has experimented successfully with remedial instruction since 1929. Under the guidance of Dr. Ernst Thelin, Chairman of the Department of Psychology, and Dr. Roland McKee, supervisor of remedial programs in the College of Liberal Arts, the author conducted a remedial reading and study aids program for twenty-one students whose gains are reported for the first semester of 1940–1941.

All students selected for the remedial program were freshmen with reading performances below the 11th centile on the reading section of Form 17, Ohio State Psychological Examination. Based on university entrance norms, the average student in the group had a reading centile of 6.

Centile ratings in Table I are based on twelfth-grade norms of the Iowa Silent Reading Test, Form Am, and represent the average senior high school centiles for the twenty-one students selected for remedial reading and study aids.

TABLE I.—AVERAGE READING CENTILES BEFORE REMEDIAL AID

IOWA SILENT READING TEST	AVERAGE CENTILE SCORE
Rate and comprehension	33.94
Rate	33.00
Comprehension	37.33
Directed reading	33.33
Poetry comprehension	36.11
Word meaning	55.94
Sentence meaning	19.50
Paragraph comprehension	
Central idea	51.66
Development of detail	25.83
Location of information	
Use of index	40.83
Selection of key words	43.88
Average total score	36.88

An average reading centile of 36.88 on the Iowa Silent Reading Test as a whole represents average tenth-grade reading performance, according to norms published in the test manual.

The Remedial Program.—The primary objective of the remedial course was to habilitate each student, if possible, to academic skills of college level. The secondary objective was to increase, if possible, the reading level of the group to college status. To what degree were these objectives attained?

Remedial Class Procedures.—During thirty-one class periods—three one-hour classes on alternate days of the week—students in remedial aids followed the general procedures outlined below:

A. GENERAL DIAGNOSTIC PROCEDURES

(1) Iowa Silent Reading Test, Form Am. (Published 1939 by World Book Company, Yonkers-on-Hudson, New York.)

(2) Analysis of skills and deficiencies in rate of reading, reading comprehension, directed reading, poetry comprehension, word meaning, sentence meaning, paragraph comprehension which includes an understanding of central ideas and the development of detail, and skills in locating information and key words.

(3) Eye examinations on the Betts' Visual Sensation and Perception Tests. (Distributed by the Keystone View Company, Meadville, Pennsylvania.)

(4) Analysis of visual examination for two-eyed efficiency, near-sightedness, far-sightedness, reading distance fusion, vertical imbalance, eye-coördination level, lateral imbalance, reading fusion at far-point, and astigmatism.

(5) Vocational interest tests, personality questionnaires, or special aptitude tests when indicated.

(6) Review of freshman health examination.

(7) Daily study schedule.

EXERCISES FOR READING IMPROVEMENT

(1) Read the summary of a chapter or an article before you read more thoroughly for detailed information.

(2) Preview a book or chapter by reading rapidly through chapter, section, and paragraph headings, bold face type, tables, pictures, etc., before you read more carefully for detail.

(3) Read assignments under a time-pressure or try to complete a given amount of material within a set limit of time.

(4) Read by the "multiple-active" method instead of the passive method; for example, plan an hour of study to allow you to glance through the entire chapter for the first five minutes, study more

XUM

carefully for the next twenty-five minutes, attempt to answer questions or recite for the next ten minutes, re-read or study material not well known or organized for the next ten minutes, recite for the next five minutes, and then make a complete re-reading of the important points for the last five minutes.

(5) Study especially the highest and lowest points of a graph, table, or chart to note in particular trends represented.

(6) Change your reading speed and methods of study to your purpose, to the difficulty or nature of the material, and to the problem at hand—in accordance with what your instructor wants you to know, the type of test you will be given, and the difficulty of the assignment.

(7) Use a different method of reading when you read to remember facts than when you read to think about a problem, or when you read to apply your information to your own problems.

(8) Read for ideas instead of words.

EXERCISES FOR WORD SKILLS

(1) Select from your textbook for flash card drills words you do not instantly recognize.

(2) Select from your textbook for flash card drills words unfamiliar and non-meaningful to you with accompanying dictionary meanings, phonetic analyses, configurational analyses, contextual analyses, or word analyses.

(3) Learn the meanings of unfamiliar or non-meaningful words you meet in your studies by using the dictionary or by noting definitions in context.

(4) Make a notebook of prefixes, suffixes, and word stems.

EXERCISES IN NOTE TAKING AND OUTLINING

(1) Take notes in your regular lecture courses.

(2) Make a brief outline of materials you study.

(3) Outline in your own words rather than in the words of the author.

(4) Paraphrase sentences, paragraphs, or chapters for clearer meanings by putting the author's thoughts into your own words.

(5) Use indention, underlining, and signs of logical rank in outlines so the important meanings are readily recognizable.

(6) Spend time thinking and evaluating lectures or books so you outline only important or critical material.

(7) Have a regular, full-sized notebook, divided into separate sections for your various school subjects, for class notes.

(8) Make your notes neat and readable.

EXERCISES FOR HABITS OF STUDY AND APPLICATION

(1) Practice the recitation method of study; that is, after reading a chapter or an assignment attempt to write a skeleton outline, recite orally to yourself or someone else, or test yourself by a true-false or essay test—and then re-study to clarify or gain additional meanings.

(2) Review when forgetting is greatest; that is, review assignments the day after they were studied, make brief weekly or semi-weekly reviews of materials covered up to the date of review.

(3) Review lessons frequently—after class, before class, weekly or semi-weekly.

(4) Prepare true-false or essay examinations for self-testing and study of assignments for overlearning.

(5) Underline, check, or in other ways indicate for later reference the most important words or sentences in your books.

(6) Diagram or make charts to summarize, organize, or make the subject-matter studied more meaningful.

(7) Direct your study by setting goals for yourself—reading for something—to outline, to recite, to take a test, etc.

(8) Study in a work situation—where lighting, heating and study materials are conducive to study and where you do your best work.

(9) Work where you will have a minimum of social distractions during your study hours.

(10) Organize your study or review material in abbreviated sentence outline form.

(11) Adapt your learning to the way in which you will use your knowledge in the future—by answering questions, reciting orally, writing a summary, solving related problems, or applying the material to life situations.

(12) Study especially for the recognition of detail when preparing for a true-false examination.

(13) Study especially for organization and recall or for more general information in preparing for essay examinations.

(14) Practice examinations constructed by yourself under time-pressure before taking time-limit tests given by your instructors.

(15) Be sure you have your notebook, pen, textbook, and other needed materials when you need them in class.

XUM

(16) Schedule your time so your assignments are always in the hands of your instructor when due.

(17) Try to memorize ideas rather than the exact words of the textbook in preparing assignments or in studying for an examination.

(18) Analyze your class work or examination papers to discover your weaknesses.

GENERAL EXERCISES IN THE REMEDIAL AID PROGRAM

(1) Keep a graphic record of reading speed and comprehension scores from textbook exercises in class or home drills.

(2) Practice three hours or more in reading phrases on a "home-made" metronoscope similar to the one demonstrated in class.

(3) Study two or more books or pamphlets on "How To Study" or "How to Increase your Reading Ability."

(4) Read for at least three hours per week simple background material in a high-school textbook of political science, economics, geography, history, English, chemistry, physics, or some other textbook in the field of your special difficulty.

(5) Read for at least three hours per week background material in digests, magazines, newspapers, or other easy reading material.

(6) Practice phrasing and speed in easy reading materials such as novels, adventure, or other fiction magazines.

(7) Keep in close contact with your instructors in the matter of your scholastic progress—in coming examinations, in make-up work, in additional aids for study in their subject fields.

(8) Face emotional or social problems of adjustment to college by gaining help from your instructors or advisors when necessary.

The reader will note that remedial exercises were designed primarily for immediate improvement in each student's course work rather than in ultimate reading test gains. Consequently remedial exercises involved individual study in the students' textbook materials and lecture notes in place of traditional remedial reading exercises directed toward improved test results.

From the remedial procedures outlined and in view of thirty-one hours of a non-credit, non-compulsory, no-grade course with a minimum of outside class drill, what gains can the psychologist expect from freshmen students in the lowest decile of the Ohio State Psychological Examination and already engaged in a full college program?

Table II presents the pre-course reading test results obtained October 17, 1940, from the Iowa Silent Reading Test, Form Am,

and the post-course reading test results obtained January 9, 1941, from the Iowa Silent Reading Test, Form Bm.

TABLE II.—AVERAGE READING CENTILES SHOWING GAINS FROM REMEDIAL AIDS

Iowa Silent Reading Test	Pre-aid	Post-aid	Gains
Rate and comprehension	33.94	75.33	41.39
Rate	33.00	68.83	35.83
Comprehension	37.33	80.88	43.55
Directed reading	33.33	71.27	37.94
Poetry comprehension	36.11	77.22	41.11
Word meaning	55.94	76.05	20.11
Sentence meaning	19.50	47.38	27.88
Paragraph comprehension			
Central idea	51.66	67.27	15.61
Development of detail	25.83	65.05	39.22
Location of information			
Use of index	40.83	73.77	32.94
Selection of key words	43.88	76.00	32.12
Average total score	36.88	73.44	36.56

An average centile of 73.44 indicates that at the completion of the remedial course the typical student was reading as efficiently or better than seventy-three per cent of twelfth-grade students, or as well as average thirteenth-grade students. First semester grades in all university courses for students in the remedial program are now being compared with grades received by students (matched by class, college, sex, intelligence, and reading ability) who did not have the advantages of a remedial reading course.

Even though an evaluation in terms of scholarship is not available at the present time for this particular study, previous investigations by Dr. Ernst Thelin, Dr. Roland McKee, and Mr. William Madden (*Evaluation of an Academic Methods Course at Syracuse University*, unpublished Master's Thesis, Syracuse University Library, 1940) show positively the superior scholastic achievements of students who have had remedial work when compared with students of similar abilities at the time of enrollment in the University who did not have remedial advantages.

Nevertheless, in the present study initial test results were average for tenth-grade students while final test results were average for thirteenth-grade students. The development of three grades in

reading skill (as measured by standardized reading tests) in the period of less than one full semester is exceptionally promising for remedial programs of the future.

Summary.—(1) Upon entering the remedial aids course, the average student was reading on a tenth-grade level, as measured by standardized reading tests.

(2) Remedial work was designed primarily for immediate improvement in the students' course work rather than in ultimate reading test gains. Consequently the remedial exercises involved study in textbook materials on individual bases in place of the usual remedial reading exercises aimed directly at improved test results.

(3) The remedial aids course was a no-grade, non-compulsory, non-credit course, and students enrolled (despite their low scholastic aptitude scores) were carrying a full fifteen-hour schedule. (This phase of the program is not recommended.)

(4) After thirty-one hours of testing and remedial work, the average student in the class was reading more efficiently than seventy-three per cent of twelfth-grade students and average or above for thirteenth-grade students. In terms of reading skills, the mean gain of thirty-six centiles represents a growth of three or more grade-years in school, as measured by standardized reading tests.

XUM

THE UTILITY OF THE INSTRUCTOR'S RATING SCALE

L. C. STECKLE

Denison University

At present we are much interested in problems of teacher-training and teaching effectiveness. This is as it should be. We still are faced, however, with the task of spanning the hiatus between talking and doing. The writer believes that the application of a well-known method of measurement effective in this area largely has been neglected.

Considered in terms of absolute accumulation, a rather respectable amount of data has been gathered. These data have, to large extent, been summarized and interpreted by Cole.[5] However, when comparisons are drawn between the studies made in this field and those having to do with traditional research in psychology, the emphasis is trifling indeed.

For the past fifteen years, more or less, articles have appeared in the literature on the construction, reliability and validity of the teacher's rating scale. From such reports it is patent that the usual academic objections to student ratings of instructors are totally without objective foundation. The student does not rate an instructor in terms of the grade received in the course. Remmers[10] found the correlation between ratings and marks to be .070. Blum[3] concludes that whatever grade a student may be attaining in a course, his estimation of the instructor's ability remains relatively constant and agrees with the average rating given this instructor by the entire student group involved.

Neither is it true that the student is an incompetent judge. Guthrie[7] found the agreement among students concerning the abilities of their teachers to be expressed by a coefficient of correlation of +.79. Further, these judgments are stable. Yearly ratings correlated +.89. It has also been shown that student judgments agree reasonably well ($r = +.72 - +.89$) with such criteria of teaching effectiveness as colleague and administrative ratings.[6]

Whether we like it or not, the only real criterion of teaching ability resides in the reaction of the student himself. Guthrie has made the point well: "Students have an opportunity for observing the quality of teaching that no fellow-teacher, head of department or school authority ever enjoys. They alone have a direct classroom acquaintance with their teachers." (7, p. 175) More than this, students can and do discriminate between traits associated with good

and poor teaching. Stalnaker and Remmers[12] found the reliability of judgment to be expressed by a correlation of above +.90. There was also no definite "halo" effect. From evidence such as this we can but conclude that if any known measure of teaching effectiveness is valid, student ratings of the ability of the instructor are equally so.

That teaching excellence *per se* is put to the test by the instructor's rating scale has been shown by Heilman and Armentrout.[8] Such factors as class size, Bernreuter scores, severity of grading, student interest, sex of instructor and maturity of rater had no demonstrable effect upon the ratings.

The value of this measuring device becomes yet more evident in the light of the recent study of Mallory, Huggins and Steinberg.[9] This novel investigation indicated that students find the most satisfaction, the most interest and do the greatest amount of work in those courses they have selected because of "preference for work with a particular professor." Such courses were also rated consistently highest in "value" and "enjoyment" over courses selected for all other reasons.

There can be no doubt that the instructor's rating scale measures that which it is designed to measure. Further, it measures this well. The obvious next step, however, has not been taken with any degree of unanimity. If the device is a valid tool, then its continued use should make for increasing efficiency of instruction. There is some evidence that it does.[1,2,4,11]

The study to be reported here shows the change in student reaction to the writer as instructor over a two-year period involving four sections of Elementary Psychology. The students, one hundred forty-eight in number, were made up principally of sophomores with a few juniors and seniors and a sprinkling of freshmen. Each student at the close of the semester anonymously rated the instructor on the following scale.

You are going to be asked to mark a series of statements as either true or false depending upon your personal opinion for the course in elementary psychology that you were registered in during this semester. This information is of interest to the administration and has no bearing on the grade that you will receive in this course or in any other course. *Please do not write your name on these papers.* The following is a sample item: T (F) This course was too difficult.*

* Constructed by F. C. Dockerry and W. L. Valentine and used at the Ohio State University. Reproduced by permission.

T F 1. The instructor appears to have command of the classroom situation.
T F 2. The instructor gives the impression of being well grounded in his field.
T F 3. The instructor frequently rants and declaims.
T F 4. The instructor fails to command the respect of his class.
T F 5. The instructor is listless and weak.
T F 6. The instructor antagonizes his students.
T F 7. The instructor is energetic and forceful.
T F 8. The instructor is willing to permit students to present their views even though they differ from him.
T F 9. The instructor has succeeded in gaining the coöperation of his students.
T F 10. The instructor pays so little attention to discipline that the class is generally in disorder.
T F 11. The instructor is ill mannered.
T F 12. The instructor displays sympathy and understanding in dealing with students.
T F 13. The instructor's presentation is so indirect that it is almost incoherent.
T F 14. The instructor has developed a friendly attitude toward students.
T F 15. The instructor frequently comes to class poorly prepared.
T F 16. The instructor is logical in his presentation.
T F 17. The instructor is fair in dealing with students.

INSTRUCTOR'S RATING SCALE*

Indicate by an X the Instructor's Position on This Scale for Each Item

	Poor			Fair			Average			Good			Excellent		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Self-control															
2. Enthusiasm															
3. Good judgment															
4. Magnetism†															
5. Considerateness															
6. Adaptability															
7. Breadth of interest															
8. Honesty															
9. Refinement															
10. Leadership															

* From Pressey, S. L. and Troyer, M. E.: *Laboratory Workbook in Applied Educational Psychology*. Harper & Brothers, p. 21. Reproduced by permission of Harper and Brothers.

† Approachability, cheerfulness, optimism, sociability.

T F 18. The instructor is available for conferences.
T F 19. The instructor failed to admit the limitations of his knowledge (bluffed).
T F 20. The instructor's expression is crude and often ungrammatical.
T F 21. The instructor avoids all difficult questions which students bring up in classes.
T F 22. The instructor appears to be well read in subjects other than that which he teaches.

Composite ratings for the school years ending in 1940 and 1941 appear in Table I.

TABLE I*

Statements			Rating scale		
Item number	1940 per cent true	1941 per cent true	Item number	Average 1940	Rating 1941
1	94	100	1	10.9	13.4
2	98	100	2	11.1	12.3
3	18	6	3	10.6	12.6
4	12	2	4	11.7	13.4
5	2	2	5	8.9	11.7
6	43	9	6	11.7	13.3
7	89	95	7	11.7	12.6
8	91	99	8	13.8	14.3
9	85	97	9	11.3	13.7
10	7	5	10	11.9	13.1
11	14	5			
12	45	81			
13	12	4			
14	89	96			
15	9	4			
16	96	94			
17	80	92			
18	100	99			
19	5	4			
20	14	2			
21	5	2			
22	87	91			

* A difference of ten percentage points under "Statements" and one of two and five-tenths points under "Rating Scale" is statistically significant.

Manifestly an instructor can improve. In order to do this, however, some determination of the *status quo* is necessary. The instruc-

tor's rating scale provides just such a measure. Further, it permits for continual check on whatever differences may appear in student reaction as a consequence of changes in teaching techniques, emphases, attitudes, etc. If the instructor can survive the blow to ego that most initial rating scale results evoke, he becomes in a position more closely to approximate the ultimate goal in teaching; that of "making the material functional in the lives of the students."

REFERENCES

(1) Anonymous: "Student Help in the Growth of a Teacher." *J. Educ. Method*, Vol. x, 1930, pp. 30–33.

(2) Armentrout, W. D.: "Improving College Teaching by Consulting the Consumers." *Sch. Exec. Mag.*, Vol. li, 1932, pp. 476–477.

(3) Blum, M. L.: "An Investigation of the Relation Between Student's Grades and their Ratings of the Instructor's Ability to Teach." *J. Educ. Psychol.*, Vol. xxvii, 1936, pp. 217–221.

(4) Clem, O. M.: "What Do My Students Think about My Teaching?" *Sch. & Soc.*, Vol. xxxi, 1930, pp. 96–100.

(5) Cole, Luella: *The Background for College Teaching*. Farrar & Rinehart, N. Y., 1940, p. 610.

(6) Douglass, H. R.: "Rating the Teaching Effectiveness of College Instructors." *Sch. & Soc.*, Vol. xxviii, 1928, pp. 192–197.

(7) Guthrie, E. R.: "Measuring Student Opinion of Teachers." *Sch. & Soc.*, Vol. xxv, 1927, pp. 175–177.

(8) Heilman, J. D., and Armentrout, W. D.: "The Rating of College Teachers on Ten Traits by their Students." *J. Educ. Psychol.*, Vol. xxvii, 1936, pp. 197–216.

(9) Mallory, E. B., Huggins, M., and Steinberg, B.: "The Relationship between Reasons Students give for Taking Certain Courses and Student Estimates of these Courses." *J. Educ. Psychol.*, Vol. xxxii, 1941, pp. 13–22.

(10) Remmers, H. H.: "The Relationship between Student Marks and Student Attitude toward Instructors." *Sch. & Soc.*, Vol. xxviii, 1928, pp. 759–760.

(11) Root, A. R.: "Student-Teacher Rapport." *J. Higher Educ.*, Vol. v, 1934, pp. 133–135.

(12) Stalnaker, J. M. and Remmers, H. H.: "Can Students Discriminate Traits Associated with Success in Teaching?" *J. App. Psychol.*, Vol. xii, 1928, pp. 602–610.

XUM

A NOTE REGARDING EXCEPTIONAL MUSICAL ABILITY IN A LOW-GRADE IMBECILE

WILLIAM A. OWENS

Co-Chairman, Division of Psychology and Education, Winona State Teachers College

AND

WALTER GRIMM

Chairman, Division of Music, Winona State Teachers College

L. has been under casual observation for several years at the Faribault, Minnesota, Colony for the Feebleminded during recurring visits to the institution. Attention was attracted by her ability to play on the piano any music hummed or sung to her regardless of the key in which it might be placed.

Inquiry and closer observation reveal the following facts about her: She is twenty-three years and three months old with a mental age of two years, nine months. Her IQ is 20.[1] She entered the institution at fourteen, in 1932. In the ward in which she lives there is a piano of poor quality on which she plays a great deal, to the enjoyment of perhaps fifty patients of approximately the same intelligence level.

They frequently sing to her playing, but neither the piano nor the voices can be thought of as refining her appreciation of pitch.

Left to herself, she plays popular music predominantly, as she hears it over the radio. She also plays hymns, many of which she hears at devotional exercises, but her response to a hummed air does not seem to depend upon familiarity. It seems rather that the amount of playing that she has done has provided an organization of responses akin to those involved in spelling a new word based on a large bulk of spelling experience involving all likely combinations.

It seemed probable, for example, that she had never played Brahms' "Lullaby," yet, with effectiveness far beyond her mental years, she played it in the key in which it was hummed to her. She has little or no ability to play a selection as named. Her responses seem to require a stimulus which she can copy with hands on the keyboard.

From the standpoint of artistry or finished performance there is naturally much to be desired. For example, her left hand is much less

[1] Case history report of state examiner on file at the institution. A retest by the state examiner (Kuhlmann Revision) on August 14, 1941 gave MA 3-7, IQ .23.

effective than her right, and its support of the right hand does not always give true harmony. However, in view of her background and intellectual level one can only be deeply impressed that her performance is as fine as it is.

Her father is a miner who, with the mother, two years younger, maintains a good home. There are four sisters. Two are now playing the piano by ear.

L's low ability was obvious before she was one year old. She was difficult to discipline both in the home and the school because of poor understanding. She entered the latter at six and was promptly excluded. She never did any useful work either at home or in the Faribault institution, because of incompetence.

The home acquired a piano when she was five or six, but she never received any instruction. No known relative has any exceptional musical ability according to report.

There is little in her history that is noteworthy. She had no conspicuous illness nor any symptoms other than those consistent with her intellectual status. She gave a positive Babinski at 14. Speech is very limited, repetitive, and indistinct.

While the kind of ability she manifests is not unknown in those of low intelligence, the degree is so far above anything we have ever observed in defectives of, perhaps, three times her mental age that it constitutes, in our thought, a noteworthy discrepancy.

XUM

PUBLICATIONS RECEIVED

CARTER ALEXANDER. *How to Locate Educational Information and Data.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 443.

CARTER ALEXANDER. *Alexander Library Experiences.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 158. (paper.)

CARTER ALEXANDER. *Instructor's Manual for use with the Alexander Library Experiences.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 15. (paper.)

F. MATTHIAS ALEXANDER. *Man's Supreme Inheritance.* New York: E. P. Dutton & Co., 1941, pp. 354.

F. MATTHIAS ALEXANDER. *The Universal Constant in Living.* New York: E. P. Dutton and Co., Inc., 1941, pp. 270.

EARL W. ANDERSON, JOHN DOUGLAS LEITH, AND RALPH F. STREBEL, Editorial Committee. *Current Practices in Institutional Teacher Placement.* Washington: National Institutional Teacher Placement Association, 1941, pp. 186.

ANDRAS ANGYAL. *Foundations for a Science of Personality.* New York: The Commonwealth Fund, 1941, pp. 398.

J. HOWELL ATWOOD, DONALD W. WYATT, VINCENT J. DAVIS, AND IRA D. WALKER. *Thus Be Their Destiny: The Personality Development of Negro Youth in Three Communities.* Washington: American Council on Education, 1941, pp. 96. (paper.)

S. HOWARD BARTLEY. *Vision: A Study of Its Basis.* New York: D. Van Nostrand Co., Inc., 1941, pp. 350.

BERNICE BAXTER. *Teacher-pupil Relationships.* New York: The Macmillan Co., 1941, pp. 166.

RUDYARD K. BENT AND HENRY H. KRONENBERG. *Principles of Secondary Education.* New York: McGraw-Hill Book Co., 1941, pp. 594.

EMMETT A. BETTS AND AGNES SUTTON AUSTIN. *Visual Problems of School Children.* Chicago: The Professional Press, Inc., 1941, pp. 80. (paper.)

ARTHUR C. BINING, WALTER H. MOHR, AND RICHARD H. MCFEELY. *Organizing the Social Studies in Secondary Schools.* New York: McGraw-Hill Book Co., Inc., 1941, pp. 337.

ARTHUR C. BINING AND DAVID H. BINING. *Teaching the Social Studies in Secondary Schools.* New York: McGraw-Hill Book Co., Inc., 1941, pp. 378.

GUY L. BOND AND EVA BOND. *Developmental Reading in High School.* New York: The Macmillan Co., 1941, pp. 363.

STEUART HENDERSON BRITT. *Social Psychology of Modern Life.* New York: Farrar and Rinehart, Inc., 1941, pp. 562.

OSCAR KRISEN BUROS. *The Second Yearbook of Research and Statistical Methodology: Books and Reviews.* Highland Park, N. J.: The Gryphon Press, 1941, pp. 383.

R. H. CHATTERTON. *Methods of Lesson Observing by Pre-service Student-teachers.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 137.

NORMA E. CUTTS AND NICHOLAS MOSELEY. *Practical School Discipline and Mental Hygiene.* Boston: Houghton Mifflin & Co., 1941, pp. 324.

EDUCATIONAL RECORDS BUREAU STAFF. 1941 *Achievement Testing Program in Independent Schools and Supplementary Studies.* New York: Educational Records Bureau, 1941, pp. 96. (paper.)

HORACE B. ENGLISH AND VICTOR RAIMY. *Studying the Individual School Child: A Manual of Guidance.* New York: Henry Holt & Co., 1941, pp. 131. (paper.)

MORDECAI EZEKIEL. *Methods of Correlation Analysis.* Second Edition. New York: John Wiley & Sons, Inc., 1941, pp. 531.

HERBERT A. FALK. *Corporal Punishment.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 162.

H. V. GARD. *Infinite Man.* Boston: Bruce Humphries, Inc., 1941, pp. 291.

CHARLES E. GERMANE AND EDITH G. GERMANE. *Personnel Work in High School.* New York: Silver, Burdett and Co., 1941, pp. 608.

WILLIAM GOLDFARB. *An Investigation of Reaction Time in Older Adults.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 76.

EDWARD B. GREENE. *Measurements of Human Behavior.* New York: The Odyssey Press, 1941, pp. 777.

GEORGE W. HARTMANN. *Educational Psychology.* New York: American Book Co., 1941, pp. 552.

S. I. HAYAKAWA. *Language in Action.* New York: Harcourt, Brace & Co., 1941, pp. 245.

SAMUEL PERKINS HAYES. *Contributions to a Psychology of Blindness.* New York: American Foundation for the Blind, Inc., 1941, pp. 296.

HARRY WALKER HEPNER. *Psychology Applied to Life and Work.* New York: Prentice-Hall, Inc., 1941, pp. 771.

High-school Reading Program, A. Maryland School Bulletin. Baltimore: Maryland State Department of Education, 1941, pp. 127.

L. THOMAS HOPKINS. *Interaction: The Democratic Process.* Boston: D. C. Heath & Co., 1941, pp. 490.

LOWRY S. HOWARD. *The Road Ahead: A College Orientation and Guidance Book.* Yonkers-on-Hudson, N. Y.: World Book Co., 1941, pp. 402.

PAUL B. JACOBSON AND WILLIAM C. REAVIS. *Duties of School Principals.* New York: Prentice-Hall, Inc., 1941, pp. 812.

WILLIAM HEARD KILPATRICK. *Selfhood and Civilization: A Study of the Self-other Process.* New York: The Macmillan Co., 1941, pp. 243.

WILLIAM ARLOW KILGORE. *Identification of Ability to Apply Principles of Physics.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 34.

HOWARD ANDREW KINHART. *The Effect of Supervision on High-school English.* Baltimore: Johns Hopkins Press, 1941, pp. 102. (paper.)

GEORGE C. KYTE. *The Principal at Work.* Boston: Ginn and Co., 1941, pp. 496.

D. WELTY LEFEVER, ARCHIE M. TURRELL, AND HENRY I. WEITZEL. *Principles and Techniques of Guidance.* New York: Roland Press Co., 1941, pp. 522.

EDITHA LUECKE. *Factors Related to Children's Participation in Certain Types of Home Activity.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 103.

MURDO MACKENZIE. *The Human Mind: The Organ of Thought in Function and Dysfunction.* Philadelphia: The Blakiston Co., 1941, pp. 216.

EARL SCHENCK MIERS. *Composing Sticks and Mortar Boards.* Camden, N. J.: The Haddon Craftsmen, 1941, pp. 97.

NEAL E. MILLER AND JOHN DOLLARD. *Social Learning and Imitation.* New Haven: Yale University Press, 1941, pp. 341.

PAUL R. MORT AND FRANCIS G. CORNELL. *American Schools in Transition.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 546.

CHARLES S. MYERS, FRANK N. FREEMAN, AND MORRIS S. VITELES. *Modern Psychology.* Philadelphia: University of Pennsylvania Press, 1941, pp. 48. (paper.)

R. C. OLDFIELD. *The Psychology of the Interview.* London: Methuen and Co., Ltd., 1941, pp. 144.

LOUISE PRICE. *Creative Group Work on the Campus.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 437.

MANDEL SHERMAN. *Basic Problems of Behavior.* New York: Longmans, Green & Co., 1941, pp. 440.

E. J. SPADINO. *Writing and Laterality Characteristics of Stuttering Children.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 82.

PAUL STUDENSKI AND PAUL R. MORT. *Centralized vs. Decentralized Government in Relation to Democracy.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 69. (paper.)

Studies in Arithmetic. Volume II. London: University of London Press, 1941, pp. 218.

J. EDWARD TODD. *Social Norms and the Behavior of College Students.* New York: Bureau of Publications, Teachers College, Columbia University, 1941, pp. 190.

XUM

A New and Up-to-date

COURSE OF STUDY IN ENGLISH

Covering the teaching of oral and written English, usage and technicalities, and reading (both the recreational and study types) for

PRIMARY GRADES

INTERMEDIATE GRADES

JUNIOR HIGH SCHOOL

based, as far as possible, upon the best that experimentation and expert opinion have produced. Prepared by the Supervisors and Teachers of

BALTIMORE COUNTY

under the direction of Dr. R. L. Lyman.

Subjects Covered

Reading: (Grades I to IX) Recreational—reading for pleasure, with emphasis upon excellent literature in both prose and poetry; for study—(1) silent reading to gain information and to improve one's skill in comprehension; (2) oral reading to give information to others and to improve one's ability in interpreting what is read.

Expression, oral and written: (Grades I to III) The expression in daily use by the pupils for such purposes as telling stories, talking together, answering questions, relating experiences.

Communication, oral and written: (Grades IV to IX) The composition in daily use by the pupils for such purposes as conversation, telling stories, answering questions, writing letters, giving directions, reporting experiences, participating in group discussions.

Language: (Grades I to III) Items dealing with the mechanics of oral and written expression, capitalization, and punctuation.

Spelling and Handwriting: (Grades I to III) The items dealing with techniques of study in spelling and handwriting.

Technicalities: (Grades IV to IX) Items dealing with the mechanics of communication, namely, sentence work and grammar, usage, capitalization, punctuation, and spelling.

Part One, Grades I to III. Part Two, Grades IV to VI. Part Three, Grades VII to IX.

Price: $1.42 per section postpaid. $3.83 per set postpaid. Special discounts on quantity orders.

WARWICK AND YORK

Publishers BALTIMORE, MD.

XUM

A Handbook for Student-Teachers

By JANE E. McALLISTER
and others.

The purpose of this Handbook is to aid the student-teacher early in his college career to realize the qualifications necessary for teaching, and to encourage and help him to direct all his efforts toward the goal of teaching.

$1.02 *postpaid*

WARWICK AND YORK

10 EAST CENTRE STREET BALTIMORE, MD.

EDUCATIONAL PUBLISHERS

Determinism in Education

By W. C. BAGLEY

Of this volume, which appeared fifteen years ago, Peter Sandiford in his recently published *Foundations of Educational Psychology* says:

"Bagley's 'Determinism in Education' is the best collection of data supporting the environmentalist's position that has been assembled, and should be read as an antidote against the extreme hereditarian position."

$2.40 *plus 3¢ postage in the United States; 12¢ to foreign countries.*

WARWICK AND YORK

10 EAST CENTRE STREET BALTIMORE, MD.

EDUCATIONAL PUBLISHERS

SCHOOL AND SOCIETY

Edited by WILLIAM C. BAGLEY

The only educational news-magazine of nation-wide circulation that is published every week throughout the calendar year

SCHOOL AND SOCIETY covers all major fields of educational endeavor, including colleges and universities; professional schools; public, private and parochial schools of elementary and secondary grade; and the library service.

In addition to the news of the week and accounts of educational happenings both at home and abroad, SCHOOL AND SOCIETY publishes articles of timely interest, annotated lists of all publications received, book reviews, and brief reports of educational research.

SUBSCRIPTION, $5.00 A YEAR

Members of the Society for the Advancement of Education, Inc., (a non-commercial non-profit organization) receive SCHOOL AND SOCIETY at a special rate of $3.50. Application forms for membership (which is limited to individuals) will be sent on request.

Address all correspondence to

The Society for the Advancement of Education, Inc.

425 WEST 123D STREET, NEW YORK, N. Y.

THE MAPLE PRESS COMPANY, YORK, PA.

Duke University Library
MAR 6 - 1942
Durham, N. C.

VOL. XXXII DECEMBER, 1941 No. 9

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

Title-page and Index Volume XXXII (1941) . . . . . i-viii

The Retention of Orally Presented Materials 641
EDWARD M. SPENCER

Student Evaluation of Knowing College Aptitude Test Score. 656
R. K. COMPTON

A Study of the Effect of Various Teaching Procedures upon Observed Group Attention in the Secondary School. 665
R. W. EDMISTON AND R. W. BRADDOCK

Deriving Language and Non-language Measures of Intelligence from the Kuhlmann-Anderson Intelligence Tests 673
GEORGE SPACHE

Two Test-retest Experiments on Knowledge of the English Language 681
AUDRY PRINDLE, R. R. REICHART AND H. R. LASLETT

Studies in the Psychology of Memorizing Piano Music. VII. A Comparison of Three Degrees of Overlearning. 688
GRACE RUBIN-RABSON

The Concept of "Highest Binet Attainment" Applied to Unselected Clinical Cases 697
EDITH A. DAVIS

Book Reviews 704

$6.00 per Year • Published Monthly September to May

WARWICK & YORK, INC.
BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

XUM

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

Established 1910

EDITED BY

JACK W. DUNLAP
University of Rochester

IN ASSOCIATION WITH

JOHN G. DARLEY
University of Minnesota

H. H. REMMERS
Purdue University

STEPHEN M. COREY
University of Chicago

PERCIVAL M. SYMONDS
Teachers College (Columbia University)

HAROLD E. JONES
University of California

PAUL A. WITTY
Northwestern University

H. E. BUCHHOLZ
Managing Editor

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

Manuscripts should be typed, double-spaced, and sent to Jack W. Dunlap, Catherine Strong Hall, University of Rochester, Rochester, N. Y. Return postage should be included with all unsolicited manuscripts. Books and other materials for review, and correspondence regarding editorial and business matters should be addressed to the Publishers.

THE JOURNAL is published monthly from September to May. The price per year in the United States and Pan-American countries is $6.00; $6.20 to Canada; and $6.40 to other foreign countries. Part-year subscriptions are 90 cents per issue ordered. Back volumes are $7.00 each, and back issues $1.10.

Subscribers should notify the Publishers of change of address at least four weeks in advance of publication of issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of next succeeding number.

WARWICK AND YORK *Publishers* BALTIMORE, MD.

XUM